DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service

Food and Drug Administration
8757 Georgia Avenue
Silver Spring MD 20910

DEC 12 1985

Clinical Diagnostics, Inc.
ATTN: Bruce G. Meyer, Jr.
300 E. Mineral Avenue, #6
Littleton, CO 80122

D.C. Number : K854903
Received : 12-9-85
Product : Gordon Diagnostic System
GDS Data Analysis Program

The Premarket Notification you have submitted as required under Section 510(k) of the Federal Food, Drug and Cosmetic Act for the above referenced device has been received and assigned a unique document control number (D.C. Number above). Please cite this D.C. Number in any future correspondence that relates to this submission.

We will notify you when the processing of this submission has been completed or if any additional information is required. You are required to wait ninety (90) days after the received date shown above or until receipt of a "substantially equivalent" letter before placing the product into commercial distribution. I suggest that you contact us if you have not been notified in writing at the end of this ninety (90) day period before you begin marketing your device. Written questions concerning the status of your submission should be sent to:

Food and Drug Administration
Center for Devices and
Radiological Health
Office of Device Evaluation
Document Mail Center (HFZ-401)
8757 Georgia Avenue
Silver Spring, Maryland 20910

If you have procedural or policy questions, please contact the Division of Small Manufacturers Assistance at their toll-free number (800) 638-2041 or me at (301) 427-8162.

Sincerely yours,

Robert I. Chissler
Premarket Notification Coordinator
Office of Device Evaluation
Center for Devices and
Radiological Health

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service

Food and Drug Administration
8757 Georgia Avenue
Silver Spring MD 20910

JUN 2 1986

Clinical Diagnostics, Inc.
Attn: Bruce Meyer
300 E. Mineral Avenue, #6
Littleton, Colorado 80122

Re: K854903B
Gordon Diagnostic System Model 1
Dated: April 30, 1986
Received: May 1, 1986

Dear Mr. Meyer:

We have reviewed your Section 510(k) notification of intent to market the above device and we have determined the device to be substantially equivalent to devices marketed in interstate commerce prior to May 28, 1976, the enactment date of the Medical Device Amendments. You may, therefore, market your device subject to the general controls provisions of the Federal Food, Drug, and Cosmetic Act (Act) until such time as your device has been classified under Section 513. At that time, if your device is classified into either class II (Performance Standards) or class III (Premarket Approval), it would be subject to additional controls. Please note: This action does not affect any obligation you might have under the Radiation Control for Health and Safety Act of 1968, or other Federal Laws or regulations.

General controls presently include regulations on annual registration, listing of devices, good manufacturing practice, labeling, and the misbranding and adulteration provisions of the Act. In the future, the scope of general controls may be broadened to include additional regulations.

All regulations and information on meetings of the device advisory committees, their recommendations, and the final decisions of the Food and Drug Administration (FDA) will be published in the Federal Register. We suggest you subscribe to this publication so you can convey your views to FDA if you desire and be notified of any additional requirements imposed on your device. Subscriptions may be obtained from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Such information also may be reviewed in the Dockets Management Branch (HFA-305), Food and Drug Administration, Room 4-62, 5600 Fishers Lane, Rockville, Maryland 20857.

This letter does not in any way denote official FDA approval of your device or its labeling. Any representation that creates an impression of official approval of this device because of compliance with the premarket notification regulations is misleading and constitutes misbranding. If you desire advice on the labeling for your device or other information on your responsibilities under the Act, please contact the Office of Compliance, Division of Compliance Operations (HFZ-320), 8757 Georgia Avenue, Silver Spring, Maryland 20910.

Sincerely yours,

George C Murray

George C. Murray, Ph.D.
Director
Division of Anesthesiology, Neurology, and Radiology Devices
Center for Devices and Radiological Health

DEPARTMENT OF HEALTH AND HUMAN SERVICES **Public Health Service**

Food and Drug Administration
Center for Devices and
Radiological Health
8757 Georgia Avenue
Silver Spring, MD 20910

APRIL 30, 1986

CLINICAL DIAGNOSTICS, INC.
ATTN: BRUCE MEYER
300 E. MINERAL AVENUE, #6
LITTLETON, CO 80122

Ref : K854903
Product : GORDON DIAGNOSTIC SYSTEM GDS DATA ANALYSIS PROGRAM

We are holding your above-referenced Premarket Notification (510(k)) for 30 days pending receipt of the additional information that was requested by the Office of Device Evaluation. This information should be submitted in duplicate to:

Food and Drug Administration
Center for Devices and
Radiological Health
Document Mail Center (HFZ-401)
8757 Georgia Avenue
Silver Spring, Maryland 20910

When your additional information is received by the Office of Device Evaluation the 90-day period will begin again.

If after 30 days the requested information is not received, we will stop reviewing your submission. Pursuant to 21 CFR 20.29, a copy of your 510(k) submission will remain in the Office of Device Evaluation. If you then wish to resubmit this 510(k) notification, a new number will be assigned and the 90-day time period will begin again.

If you have procedural or policy questions, please contact the Division of Small Manufacturers Assistance at their toll-free number (800) 638-2041 or me at (301) 427-8162

Sincerely yours,

Robert I. Chissler
Premarket Notification Coordinator
Office of Device Evaluation
Center for Devices and
Radiological Health

DEPARTMENT OF HEALTH AND HUMAN SERVICES **Public Health Service**

Food and Drug Administration
Center for Devices and
Radiological Health
8757 Georgia Avenue
Silver Spring, MD 20910

FEBRUARY 10, 1986

CLINICAL DIAGNOSTICS, INC.
ATTN: BRUCE MEYER
300 E. MINERAL AVENUE, #6
LITTLETON, CO 80122

Ref : K854903
Product : GORDON DIAGNOSTIC SYSTEM GDS DATA ANALYSIS PROGRAM

We are holding your above-referenced Premarket Notification (510(k)) for 30 days pending receipt of the additional information that was requested by the Office of Device Evaluation. This information should be submitted in duplicate to:

Food and Drug Administration
Center for Devices and
Radiological Health
Document Mail Center (HFZ-401)
8757 Georgia Avenue
Silver Spring, Maryland 20910

When your additional information is received by the Office of Device Evaluation the 90-day period will begin again.

If after 30 days the requested information is not received, we will stop reviewing your submission. Pursuant to 21 CFR 20.29, a copy of your 510(k) submission will remain in the Office of Device Evaluation. If you then wish to resubmit this 510(k) notification, a new number will be assigned and the 90-day time period will begin again.

If you have procedural or policy questions, please contact the Division of Small Manufacturers Assistance at their toll-free number (800) 638-2041 or me at (301) 427-8162

Sincerely yours,

Robert I. Chissler
Premarket Notification Coordinator
Office of Device Evaluation
Center for Devices and
Radiological Health

Que K854903 84 1 1

IMAGE EVALUATION TEST TARGET (MT-3)

150mm

6″

PHOTOGRAPHIC SCIENCES CORPORATION
770 BASKET ROAD
P.O. BOX 338
WEBSTER, NEW YORK 14580
(716) 265-1600

K854903

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service

Food and Drug Administration
8757 Georgia Avenue
Silver Spring MD 20910

JUN 2 1986

Clinical Diagnostics, Inc.
Attn: Bruce Meyer
300 E. Mineral Avenue, #6
Littleton, Colorado 80122

Re: K854903B
Gordon Diagnostic System Model 1
Dated: April 30, 1986
Received: May 1, 1986

Dear Mr. Meyer:

We have reviewed your Section 510(k) notification of intent to market the above device and we have determined the device to be substantially equivalent to devices marketed in interstate commerce prior to May 28, 1976, the enactment date of the Medical Device Amendments. You may, therefore, market your device subject to the general controls provisions of the Federal Food, Drug, and Cosmetic Act (Act) until such time as your device has been classified under Section 513. At that time, if your device is classified into either class II (Performance Standards) or class III (Premarket Approval), it would be subject to additional controls. Please note: This action does not affect any obligation you might have under the Radiation Control for Health and Safety Act of 1968, or other Federal Laws or regulations.

General controls presently include regulations on annual registration, listing of devices, good manufacturing practice, labeling, and the misbranding and adulteration provisions of the Act. In the future, the scope of general controls may be broadened to include additional regulations.

All regulations and information on meetings of the device advisory committees, their recommendations, and the final decisions of the Food and Drug Administration (FDA) will be published in the Federal Register. We suggest you subscribe to this publication so you can convey your views to FDA if you desire and be notified of any additional requirements imposed on your device. Subscriptions may be obtained from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Such information also may be reviewed in the Dockets Management Branch (HFA-305), Food and Drug Administration, Room 4-62, 5600 Fishers Lane, Rockville, Maryland 20857.

This letter does not in any way denote official FDA approval of your device or its labeling. Any representation that creates an impression of official approval of this device because of compliance with the premarket notification regulations is misleading and constitutes misbranding. If you desire advice on the labeling for your device or other information on your responsibilities under the Act, please contact the Office of Compliance, Division of Compliance Operations (HFZ-320), 8757 Georgia Avenue, Silver Spring, Maryland 20910.

Sincerely yours,

George C Murray

George C. Murray, Ph.D.
Director
Division of Anesthesiology, Neurology, and Radiology Devices
Center for Devices and Radiological Health

DEPARTMENT OF HEALTH & HUMAN SERVICES — Public Health Service

Memorandum

Date: May 29, 1986

From: REVIEWER(S) - NAME(S) R.F. Munzner, Ph.D.

Subject: 510(k) NOTIFICATION K854903/B

To: THE RECORD

It is my recommendation that the subject 510(k) Notification:

✓ (A) Is substantially equivalent to marketed devices.

____ (B) Requires premarket approval. NOT substantially equivalent to marketed devices.

____ (C) Requires more data.

____ (D) Is an incomplete submission. (See Submission Sheet).

Additional Comments:

The submitter requests:

✓ No Confidentiality

____ Confidentiality for 90 days

____ Continued Confidentiality exceeding 90 days

Class Code w/Panel: 84LQD

ATTENTION TASK PERFORMANCE RECORDER

REVIEW: [signature], Ph.D. (BRANCH CHIEF) 5/30/86 (DATE)

FINAL REVIEW: [signature] (DIVISION DIRECTOR) 5/30/86 (DATE)

510(k) Review

K854903 B

Company Name CLINICAL DIAGNOSTICS, INC.

Device Name GORDON DIAGNOSTIC SYSTEM MODEL 1

	YES	NO
1. Life-supporting or life-sustaining?		✓
2. Implant (short-term or long-term)?		✓

3. Similar preenactment device(s): CONTINUOUS PERFORMANCE TEST (ROSVOLD) (device name, manufacturer) and other psychological performance tests

4. Differs from preenactment devices how?

(1) Psychological performance tests were not considered by an FDA classification panel and no classification regulation or proposal has been written.

(2) It is not clear wether these devices are medical devices or not.

(3) Labeling and intended use appear to meet the definition of a "device;" therefore, I believe the 510(k) is appropriate.

5. If appropriate: provides comparative in vitro data: see publication
provides a summary of animal testing? N/A
provides a summary of clinical testing? extensive

6. I believe this is equivalent to device(s): # none
Classification should be based on:

~~Subsection~~ UNCLASSIFIED (name) (presently Class ———).

(sign & date) [signature] 5/29/86

I believe this is not equivalent to any preenactment device.

I believe clinical testing is required before a determination can be made.

MEMORANDUM OF TELEPHONE CONVERSATION

Between: Millie Clark
Clinical Diagnostics
(303) 795-0438

And: Chief, Neurological Devices Branch, HFZ-430
Office of Device Evaluation, DANRD

Date: June 2, 1986

Subject: K854903 B

Summary:

I was told Bruce Meyer was out of town. I was refered to Millie Clark, who said she was acting on behalf of Mr. Meyer regarding this 510(k).

I pointed out that the labeling might not meet the requirements of §801.109. She agreed to do the following:

1. a precise statement of intended use would be provided under "indications". This statement would describe the device as an aid in diagnoses, and would not imply that it ~~was~~ could determine a diagnoses without other information.
2. References to specific drug regimens would be deleted from the device labeling.

Robert P. Munzner, Ph.D.

cc:

DEPARTMENT OF HEALTH AND HUMAN SERVICES **Public Health Service**

Food and Drug Administration
Center for Devices and
Radiological Health
8757 Georgia Avenue
Silver Spring, MD 20910

MAY 1, 1986

CLINICAL DIAGNOSTICS, INC.
ATTN: BRUCE MEYER
300 E. MINERAL AVENUE, #6
LITTLETON, CO 80122

D.C. Number : K854903
Received : 05-01-86
Product : GORDON DIAGNOSTIC SYSTEM GDS DATA ANALYSIS PROGRAM

The additional information you have submitted has been received.

-- We will notify you when the processing of your submission has been completed or if any additional information is required. You are required to wait ninety (90) days after the received date shown above or until receipt of a "substantially equivalent" letter before placing the product into commercial distribution. I suggest that you contact us if you have not been notified in writing at the end of this ninety (90) day period before you begin marketing you device. Written questions concerning the status of your submission should be sent to:

Food and Drug Administration
Center for Devices and
Radiological Health
Office of Device Evaluation
Document Mail Center (HFZ-401)
8757 Georgia Avenue
Silver Spring, Maryland 20910

If you have procedural or policy questions, please contact the Division of Small Manufacturers Assistance at their toll-free number (800) 638-2041 or me at (301) 427-8162

Sincerely yours,

Robert I. Chissler
Premarket Notification Coordinator
Office of Device Evaluation
Center for Devices and
Radiological Health

STATE UNIVERSITY OF NEW YORK

UPSTATE MEDICAL CENTER
750 EAST ADAMS STREET
SYRACUSE, NEW YORK 13210

K854903/B

COLLEGE OF MEDICINE
DEPARTMENT OF PSYCHIATRY AND BEHAVIORAL SCIENCES

AREA CODE 315
473-8100

April 30, 1986

Food and Drug Administration
Center for Devices and Radiological Health
Office of Device Evaluation
Document Mail Center
(HFZ - 401)
8757 Georgia Avenue
Silver Spring, Maryland 20910

DC#: K854903
Product: Gordon Diagnostic System

To whom it may concern:

Enclosed please find materials requested by Mr. Stephen Hinckley regarding our submission for FDA approval.

I would appreciate a rapid forwarding of these documents to Mr. Hinckley.

Sincerely,

Michael Gordon

Michael Gordon, Ph.D.
Associate Professor of Psychiatry

RECEIVED
1986 MAY -1 AM 11:12
DOCUMENT CONTROL CENTER

STATE UNIVERSITY OF NEW YORK

UPSTATE MEDICAL CENTER
750 EAST ADAMS STREET
SYRACUSE, NEW YORK 13210

COLLEGE OF MEDICINE
DEPARTMENT OF PSYCHIATRY AND BEHAVIORAL SCIENCES

AREA CODE 315
473-8100

April 30, 1986

Mr. Stephen Hinckley
Food and Drug Administration
8757 Georgia Avenue
Silver Spring, MD 20910

Dear Mr. Hinckley:

Enclosed please find a series of reprints regarding the various issues we discussed today by phone. You will note that I have highlighted sections that would seem most relevant to the equivalency issue. Also included are instructions manuals and advertisements to similar programs that are marketed for clinical use. Perhaps the most current and useful information for your purposes is found in the short reviews from Psychopharmacology Bulletin.

The articles by Rapport and Barkely, cited in an earlier response to the FDA review, are also included. Finally, there is a basic description of the standardization process and a copy of a page from a widely-used psychological test, which demonstrates the conventionality of our statistical approach to presenting normative data.

I hope that, in the rush to send these materials to you, we have not forgotten anything essential. I have tried to be somewhat conservative in selecting articles so that that the approval process does not become further delayed by a bad case of eye strain.

I enjoyed talking with you and trust you will contact me with any further questions or comments.

Sincerely,

Michael Gordon

Michael Gordon, Ph.D.
Associate Professor of Psychiatry

Vigilance Performance
Attention Response task recorder
attention Response Task Recorder
attention Task Response Demands

ARTICLES RELEVANT
TO EQUIVALENCY
ISSUE

ORIGINAL CPT ARTICLE

Journal of Consulting Psychology
Vol. 20, No. 5, 1956

A Continuous Performance Test of Brain Damage[1,2]

H. Enger Rosvold,[3] Allan F. Mirsky,[3] Irwin Sarason,[4]
Edwin D. Bransome, Jr.,[5] and Lloyd H. Beck[6]
Yale University

This paper presents the results of an investigation using a new instrument for the study of brain damage in human subjects. The design of the instrument, the Continuous Performance Test (CPT) was based on certain electroencephalographic evidence which suggested that brain-damaged individuals should show inferior ability as compared with non-brain-damaged individuals on tasks requiring sustained attention or alertness.

The waking EEGs of brain-damaged patients generally show either random bursts of hypersynchronous (high amplitude) activity intruding upon the normal activity of the brain, or a general hypersynchrony (3, 8). Hypersynchronous activity is also evident in the recording from the brain of a sleeping subject (3). If hypersynchrony is associated with reduced vigilance or attention, as suggested by its presence during sleep, then the hypersynchrony of the brain-damaged patients might also indicate reduced attention. For example, according to Hebb (1), hypersynchrony might interfere with the sequential firing of cell assemblies and thereby disrupt the process of attention. Other research strongly suggests that some such relationship does in fact exist. Thus, on the basis of a study of 75 prefrontal lobotomy cases, Levin states:

> Immediately following operation the patient usually lapsed into a drowsy-anergic state from which he could be temporarily aroused by stimuli, only to lapse once again into somnolence when the stimuli were removed. During this drowsy-anergic state EEGs showed diffuse slow activity and a good deal of baseline oscillations of ½ to 2 per second anteriorly mixed with slow rhythms of 2 to 6 per second frequency. As the patient emerged from the drowsy stage the preoperative pattern usually asserted itself (2, p. 422).

The usual measures of attention or alertness such as the digit span and digit symbol substitution subtests of the Wechsler-Bellevue have not consistently showed decline following brain damage (4, 5, 6, 7). In the case of the patient who shows intermittent bursts of hypersynchronous activity, and hence only momentary lapses in attention, the test performance may be due to the fact that the *S* can to a great extent choose his own time to respond, and may reorganize his attention between momentary lapses. The lapses would then not affect his score to any measurable extent.

In the case of the individual who shows a generalized hypersynchrony of the electrical

[1] The authors would like to extend their appreciation to the following individuals and institutions for providing advice and/or subjects and testing facilities for this research: Mr. Samuel Greenhouse of the National Institute of Mental Health, Bethesda, Maryland; Dr. C. Edward Stull and the Southbury Training School, Southbury, Connecticut; Drs. Walter Landmesser and Russell Fuldner and the Newington Crippled Children's Hospital, Newington, Connecticut; Dr. Rhoades and the New Haven Cerebral Palsy Clinic, New Haven, Connecticut; Drs. Milton Senn, William German, Janice Stevens, Morris Wessell, Ethelyn Klatskin, and the staff of the several pediatric and neurological clinics of the Grace-New Haven Community Hospital, New Haven, Connecticut.

[2] Supported in part by grants to H. Enger Rosvold from the Veteran's Administration, Contract VA1001-M3222 and the National Science Foundation.

[3] Now at the National Institute of Mental Health, Bethesda, Maryland.

[4] Now at the West Haven V.A. Hospital, West Haven, Connecticut.

[5] Now at the College of Physicians and Surgeons, Columbia University, New York, N. Y.

[6] Now at the University of Michigan, Ann Arbor, Michigan.

activity of the brain, the classical tests might not reflect his deficiency because they do not require sustained attention over a sufficiently long period of time.

These considerations suggested that a test which would not allow the patient to choose his own time to respond, and which would require a high level of continuous attention over an appreciable interval of time might reflect a deficit that other procedures would miss. Accordingly, the CPT was designed to provide two attention tasks, labeled *X* and *AX*, with *AX* designed to be the more difficult. This paper presents data on the CPT performance of comparable groups of brain-damaged and non-brain-damaged individuals.

If the research demonstrates that there are indeed differences between brain-damaged and non-brain-damaged groups on this test, one of the questions which arises is whether or not the differences are large enough to be of diagnostic use. The data also provide a preliminary answer to this question.

Method

Subjects. No attempt was made in this study to select patients on the basis of either locus, extent, or type of brain damage. The term "brain-damaged" is here used to include any individual for whom there was medical evidence (surgical notes, neurological examination, EEG) of brain pathology. Three groups were studied, each consisting of a brain-damaged and a non-brain-damaged subgroup. The *Defective* group consisted of 72 institutionalized feebleminded *S*s, who were diagnosed as either of "organic etiology" (the *Brain-Damaged-Defective* subgroup) or of "familial or idiopathic etiology" (the *Non-Brain-Damaged-Defective* subgroup). The *Child* group comprised 45 children from a children's hospital and from various pediatric clinics who were being treated for brain disorders such as cerebral palsy (the *Brain-Damaged-Child* subgroup) or for non-nervous-system disorders (the *Non-Brain-Damaged-Child* subgroup). The third group, designated the Adult group, consisted of 50 adult *S*s. Half of these (the *Brain-Damaged-Adult* subgroup) were either epileptics being treated at a seizure clinic or brain surgery patients being seen at a neurosurgery clinic. The remaining 24 *S*s (the *Non-Brain-Damaged-Adult* subgroup) were either part of the general medical population of a hospital or were nonpatient *S*s.

The age and a measure of intelligence were obtained for all *S*s. In addition, the length of institutionalization was obtained for the Defective group. The *S*s in the Defective group were routinely tested on the Stanford-Binet upon admission to the training school; all other *S*s were tested for IQ just prior to presenting them with the CPT. The intelligence measure used for the Child group was the Stanford-Binet vocabulary, while the *S*s in the Adult group were tested on either the Full Scale Wechsler-Bellevue, the Verbal Scale of the Wechsler-Bellevue, or the 20-minute modification of the Otis Self-administering Test of Mental Ability, Form B. Approximately equal numbers of *S*s in each of the Adult subgroups were tested on each of the three measures of intelligence. The average ages and IQs of the several groups are presented in Table 1. The mean age of the Non-Brain-Damaged-Child subgroup was significantly higher than that of the Brain-Damaged-

Table 1

Mean Ages and IQs of the Subjects

Group	*N*	Mean age	*SD* age	Mean IQ	*SD* IQ
Defectives					
Brain-damaged	29	27.3	8.1	63.3	7.8
Non-brain-damaged	43	25.1	7.8	66.8	7.5
Children					
Brain-damaged	19	9.0**	3.0	101.9	28.2
Non-brain-damaged	26	12.3**	2.9	104.6	21.2
Adults					
Brain-damaged	25	31.7	12.1	90.5*	15.2
Non-brain-damaged	25	31.5	13.9	100.0*	15.0

* Indicates a significant difference ($p < .05$).
** Indicates a significant difference ($p < .01$).

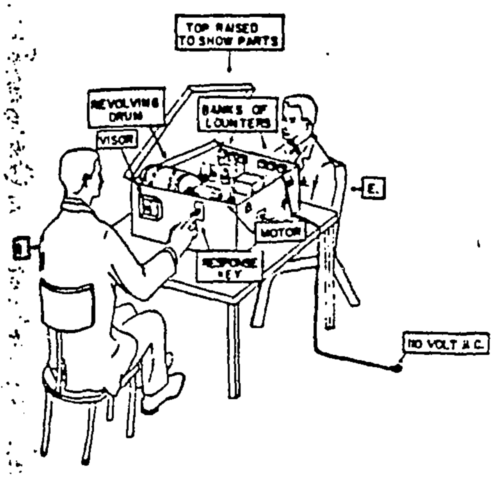


Fig. 1. Semischematic diagram showing the Continuous Performance Test apparatus, a subject and examiner in the positions they would occupy during a test.

Child subgroup ($p < .01$). Also, the IQ of the Non-Brain-Damaged-Adult subgroup was significantly higher than that of the Brain-Damaged-Adult subgroup ($p < .05$). There were no other significant differences between the various paired subgroups with respect to age or IQ. The mean length of institutionalization for the Brain-Damaged-Defective subgroup was 11.4 years, for the Non-Brain-Damaged-Defective subgroup, 9.3 years. This difference was not significant.

Apparatus. The testing instrument is illustrated in Fig. 1. It consisted essentially of a revolving drum on which two series, each of 31 letters, were mounted side by side.

One series served for the *X* task, the other for *AX* task. The drum revolved slowly, approximately twice a minute, in a boxlike case equipped with a visor through which the *S* was required to look. The *E* shifted the task from *X* to *AX* by changing the horizontal position of the visor. The letters were illuminated briefly one at a time. The illumination was provided by a bank of five 0.5-watt neon bulbs mounted inside the case, fired by means of the discharging of an 8-microfarad condenser. The *S* performed by pressing a response key when certain "correct" letters appeared. In the *X* task, every *X* was a correct letter; in the *AX* task, every *X* following an *A* was a correct letter. The letters came into view at approximately 0.92-sec. intervals. A response was scored correct if the response key was pressed within 0.69 seconds after the brief illumination of the correct letter. Following this interval, there was a period of approximately 0.23 seconds before the next letter was illuminated. Correct and incorrect responses were recorded automatically on banks of counters facing the back of the machine where *E* sat. The *X* task was always presented first and included eight *X*'s in the series of thirty-one letters. The responses occurring over the first two revolutions of the drum served as practice and were not counted in the *S*'s score. The *Ss* were then given either a 5-minute test (10 revolutions) in which there were 80 possible correct *X* responses, or a 10-minute test (20 revolutions) in which there were 160 possible correct responses. Following a 2-minute rest period, the visor was moved by *E* so that the *AX* series was in view. There were six correct *AX* sequences among the thirty-one letters of the *AX* series. As with the *X* task, the responses occurring during the first two drum revolutions served as practice. Then, the *Ss* received either a 5-minute test in which there were 60 possible correct *AX* responses, or a 10-minute test, in which there were 120 possible correct.

Procedure. No attempt was made to keep the instructions completely identical from *S* to *S;* the typical form, however, did not vary greatly from this content:

"Do you see this viewer? When the machine starts, you will see letters appearing one at a time. Your job is to press this key every time you see an *X*. Don't press it for any other letter and always press it when you see an *X*. Press the key down for a bit when you see an *X;* don't hit it too quickly. Try holding your eyes different distances from the viewer until you find the one that is most comfortable for you. Remember, press the key every time you see an *X*, but not for any other letter." Those *Ss* who had difficulty in understanding were given further instruction of the same sort.

With the *X* task and the two-minute rest period completed, the *S* was instructed as follows:

"Your job in this next part is again to press the key when you see an *X*, but this time, only if the *X* follows right after an *A*. That is, when you see the series *A—X—* press the key on the *X*. Don't press the key for any other letter following an *A* except *X*, and always press it then. Remember now, when you see the letter *A*, get set; if an *X* comes right after it, press the key." Additional instruction was given to those who had difficulty in understanding.

Except for the children, all *Ss* were tested for 10 minutes on each of the two tasks. With children, the time was reduced to 5 minutes on each task, and in addition, the instructions were modified to some extent, presenting the task as a game. For the children under the age of 10, pretesting was carried out to make certain that they could discriminate letters, and to see whether they could follow the instructions by requiring answers to two special test sheets, one containing *X*s among other letters, the other, *A*s and *X*s among other letters. Those who failed this pretest were not tested further. No *S* was tested who could not perform the minimal motor tasks of holding his head up unsupported and pressing the response key. Since the testing was conducted at many different clinics and hospitals, it was not feasible to control exactly the amount of illumination in the testing room. The ordinary lighting of the room was used without any modification; care was taken however, to insure that there was no glare on the visor of the testing apparatus.

Reliability. Two measures of the reliability of the CPT were made, using *S*s of the Defective group. Odd-even reliability was determined from the scores of 21 non-brain-damaged *S*s, by comparing the subtotal of correct responses on the 10 even revolutions of the drum with the subtotal on the 10 odd revolutions. The Pearson r for the *X* task was 0.88, for the *AX* task 0.86. To determine test-retest reliability, 21 brain-damaged and 22 non-brain-damaged *S*s were retested from four to seven weeks after the initial testing. The test-retest r for the number of correct responses of the brain-damaged group on *X* was 0.90, for *AX*, 0.79. The test-retest r for the number of correct responses of the non-brain-damaged group on *X* was 0.88, for *AX*, 0.74.

CPT scores. The *X* and *AX* responses of each *S* were scored in two ways. The absolute percentage correct was defined as the number of correct responses the *S* made divided by the number of correct responses (× 100) possible in the time allowed. The relative percentage correct was defined as the number of correct responses (× 100) the *S* made divided by his number of attempts (both correct and incorrect) during the time allowed. In the event that the two scores differed markedly, depending on the *S*'s pattern of responding, the two measures provided a more complete description of his performance than would either of the two taken singly.

Results

CPT scores. Table 2 presents the means and standard deviations of the absolute percentage of possible correct responses on the *X* and *AX* tasks for the six subgroups. Since there was a significant age difference between the Child subgroups and a significant IQ difference between the Adult subgroups (Table 1), the differences between the paired subgroup means on *X* and *AX* in these two groups were evaluated by means of an analysis of covariance. The means of the Child group in Table 2 have been adjusted for regression on age and the means of the Adult group have been adjusted for regression on IQ. The covariance procedure was corrected for heterogeneity of variance between the subgroups when the analysis required it. The mean differences in the Defective group were evaluated by means of the t test.

The differences between the subgroup means were highly significant on both the *X* and

Table 2

Adjusted Means and Standard Deviations for Child and Adult Groups, Means and Standard Deviations for Defective Groups, *X* and *AX* Absolute Percentage Correct Measures

	Brain-damaged			Non-brain-damaged			
Group	*N*	Adj. mean	*SD*	*N*	Adj. mean	*SD*	p diff.
Children							
X	19	41.61	28.39	26	82.57	14.65	$<.0005$
AX	19	46.46	28.20	26	79.41	17.44	$<.0005$
Adults							
X	25	78.53	28.26	25	88.00	17.21	$>.30$
AX	23	61.40	26.66	23	78.77	15.33	$<.01$
Defectives							
X	29	46.44	32.02	43	67.86	25.62	$<.01$
AX	29	32.30	25.79	43	54.74	20.02	$<.0005$

Table 3

Significance Levels of Differences Between X and AX Scores, for Relative and Absolute Measures of Performance, All Groups

Group	Brain-damaged		Non-brain-damaged	
	Absolute	Relative	Absolute	Relative
Children	$>.70$	$>.30$	$>.50$	$>.70$
Adults	$<.001$	$<.01$	$<.01$	$<.05$
Defectives	$<.01$	$<.001$	$<.001$	$<.001$

Table 4

Significance Levels of Differences Between Adjusted AX Mean Scores of Brain-Damaged and Non-Brain-Damaged Subgroups, All Groups

Group	Performance measure	
	Absolute	Relative
Children	$>.20$	$<.05$
Adults	$<.01$	$>.05$
Defectives	$<.05$	$<.05$

AX tasks in the Defective group ($p < .01$ and $p < .005$, respectively) and also in the Child group ($p < .0005$ in each case). In the Adult group, the brain-damaged subgroup was significantly inferior to its controls on AX ($p < .01$), but not on X ($p > .30$).

When the *relative* percentage scores of the paired Child and Adult subgroups were compared by means of an analysis of covariance, and those of the Defective group by means of the t test, the paired subgroup differences were all in the same direction and of approximately the same magnitude as those reported in Table 2. The levels of significance were in no case less than the corresponding levels reported in Table 2. The only difference was that the *relative* scores on the X task discriminated significantly ($p < .05$) between the Adult subgroups, whereas the absolute scores on the X task (Table 2) did not.

In order to determine if the AX task was in fact more difficult than the X task, the mean X score was compared with the mean AX score for each of the six subgroups. Table 3 presents the significance levels from the t test of the differences between X and AX scores for both the relative and absolute measures of performance in all six subgroups. Each of the Defective and Adult subgroups performed significantly worse (at $p < .05$ or less) on the AX than on the X task. With the children, none of the differences between X and AX was significant. To determine if the brain-damaged subgroups had relatively more difficulty on AX (i.e., independent of their performance on X) than the non-brain-damaged subgroups, an analysis of covariance was run on the AX scores of the paired brain-damaged and non-brain-damaged subgroups with X scores as the covariate. Table 4, which presents the results of the analysis, shows that in terms of the absolute measures

Table 5

Percentages of Correct Identification of Diagnostic Category by CPT Scores, All Groups

Subgroup	Defectives		Children		Adults	
	Brain-damaged	Non-brain-damaged	Brain-damaged	Non-brain-damaged	Brain-damaged	Non-brain-damaged
X	29	43	19	26	25	25
$X-R$	75.9***	67.4**	84.2***	76.9***	68.0*	68.0*
$X-A$	62.0	58.1	89.5***	76.9***	64.0	64.0
$AX-R$	72.4**	65.1*	84.2***	76.9***	66.7*	66.7*
$AX-A$	75.9***	67.4**	84.2***	76.9***	65.0†	65.0†

Note: the spanning header "Group" is printed above Defectives, Children and Adults.

* $p < .05$.
** $p < .01$.
*** $p < .005$.
† Based on 23 subjects.
Note.—p-values refer to the probability that the percentage differs from chance (50%). Nonstarred percentages do not differ significantly from chance.

of performance, the Brain-Damaged-Adult and -Defective subgroups had significantly greater difficulty on the *AX* task than their respective nondamaged subgroups, even when the paired subgroups were adjusted to equal levels of performance on the *X* task. Similarly, in terms of the relative measures of performance, the Brain-Damaged-Child and -Defective subgroups had significantly greater difficulty on *AX* than their respective non-damaged subgroups.

Correct identification of subgroup. Table 5 presents the percentages of correct identification of diagnostic category of all the subgroups on all measures yielded by the CPT. These were determined for each subgroup on each CPT measure from the over-all median of scores of a group on that measure. For example, considering the *X* scores in the Defective group, the percentage of brain-damaged *Ss* scoring *below* the combined Defective group median defines the percentage of correct identification of brain-damaged defectives on the *X* task. Similarly, the percent-

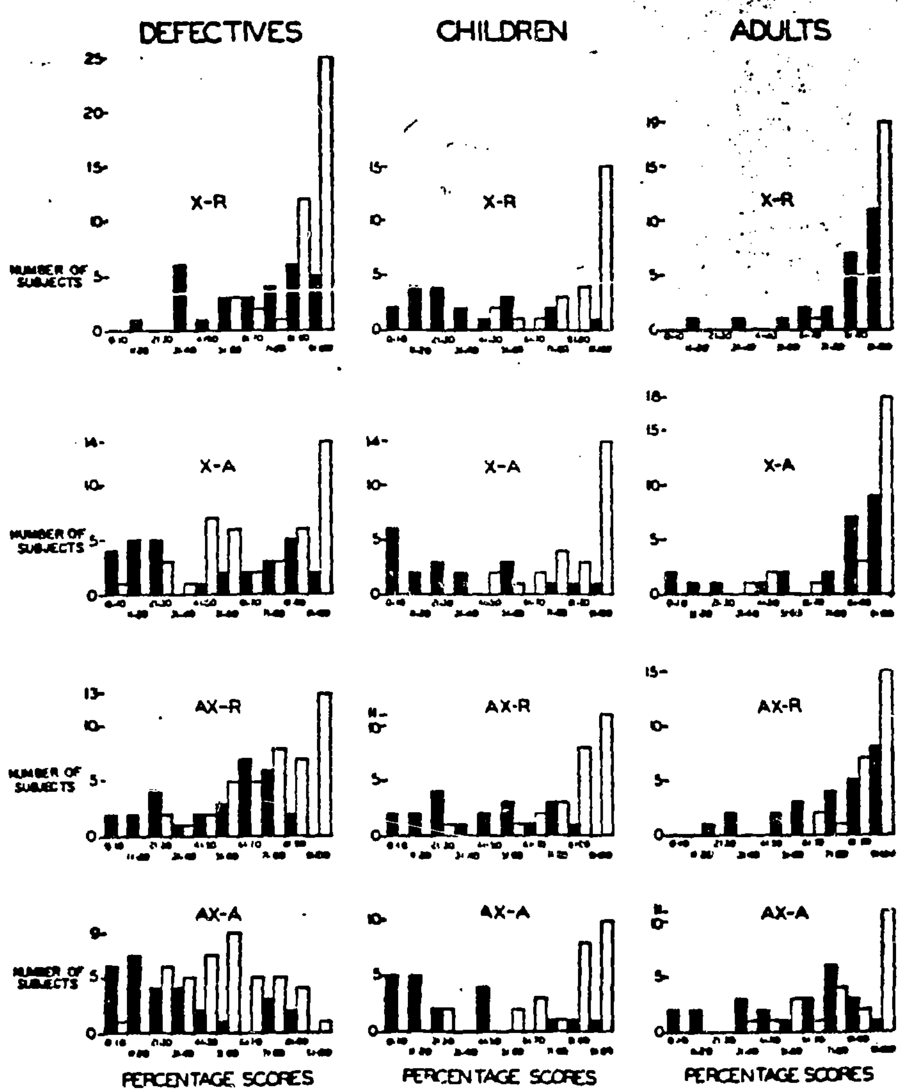


Fig. 2. Frequency distributions of subjects by percentage correct score intervals, all groups, all CPT measures. Black columns represent frequencies of brain-damaged subjects, white columns, frequencies of non-brain-damaged controls.

age of non-brain-damaged defectives scoring above the over-all median defines the percentage of correct identification of non-brain-damaged defectives on the *X* task. These percentages were computed to provide an estimate of the potential usefulness of the several scores of the CPT as diagnostic indicators in various clinical groups. In Table 5, *X-R* and *AX-R* refer to the relative measures of performance on these tasks; *X-A* and *AX-A* refer to the absolute measures of performance. The relative measures provided more significant percentages of correct identification than did the absolute measures; all of the six *X-R* and all of the six *AX-R* subgroup percentages were significant at $p < .05$. The Child group had more significant percentages of correct identification than any other (eight of eight) all of which were significant at $p < .005$. There were equal numbers of significant percentages of correct identification among the non-brain-damaged and brain-damaged subgroups; however, these percentages tended to be higher among the brain-damaged subgroups.

Figure 2 presents for each group and CPT measure, the distribution by percentage score of brain-damaged and non-brain-damaged *Ss* (uncorrected for age or IQ differences between subgroups). Figure 2 emphasizes in general the great range in scores achieved by brain-damaged *Ss* as compared with non-brain-damaged *Ss*. The scores of the non-brain-damaged *Ss* are usually concentrated near the upper end of the distribution; in fact, the modal score for non-brain-damaged *Ss* is in the 91 to 100% interval in every case except for the *AX-A* measure in the Defective group. In contrast with this, the scores of the brain-damaged *Ss* span the entire distribution, or the entire distribution excepting one score interval, in every case.

Discussion

The results of this study indicate that brain-damaged individuals perform poorly relative to non-brain-damaged controls on a task requiring continuous attention. The results also indicate that on a task designed to require even more sustained attention the brain-damaged individuals perform even more poorly. The score distributions of the various subgroups (Figure 2) and the percentages of individuals correctly identified by diagnostic category (Table 5) suggest that the CPT is sufficiently sensitive to the effects of brain damage so that it might prove useful clinically. The data of Table 5 suggest that the relative measures are more discriminating than the absolute measures, that the test is somewhat more reliable in reflecting the presence than the absence of brain damage and that it is most discriminating when applied to the children of this study.

These findings are of course preliminary. Use of the CPT in the clinic must await standardization on a variety of diagnostic groups, to ascertain whether brain pathology only and which types of brain pathology lead to deficient performance. For example, some data which have not been reported here indicate that schizophrenics perform no differently from normals of comparable age and IQ. More information concerning other groups is required. Standardization on a variety of diagnostic groups is also necessary in order to have a reliable diagnostic score for each subgroup. Thus, in the present study, cutoff points which included most brain-damaged children missed many brain-damaged adults (see Figure 2).

Diagnostic considerations aside, the question remains as to whether the impaired performance of the brain-damaged *Ss* on the CPT is in fact due to an impairment of attention, whether in the form of momentary lapses, or of a lowering of a general level. Clearly, the available information does not rule out alternate interpretations of the deficit. To make certain that it is attention that is impaired would require more extensive information about the *S*'s test behavior than the present instrument provides. It is clear, furthermore, that the available information does not permit the conclusion that the impaired CPT performance is indeed related to hypersynchrony as evident in the EEG. Nonetheless, the results of this study do lend support to these hypotheses suggesting that it might now be profitable to make a more direct test of the original formulation. This test would be provided by simultaneously measuring CPT performance and brain activity to determine whether errors on the CPT are coincident in time with hypersynchronous patterns on the EEG and whether there is a correlation between degree of deficit on the CPT and degree of hypersynchrony.

Summary

1. The Continuous Performance Test, a procedure for the detection and study of brain damage in humans, is described.

2. Three groups of *S*s, each including a brain-damaged and a non-brain-damaged subgroup, were tested on this procedure.

3. The brain-damaged subgroups were significantly inferior to their non-brain-damaged controls on the measures yielded by the CPT, and these differences were increased when the difficulty of the task was increased.

4. The CPT is sufficiently reliable and yields sufficiently large differences between subgroups to suggest that it might ultimately prove useful as a clinical instrument for the diagnosis of brain damage.

5. An interpretation of the inferior performance of the brain-damaged *S*s was offered in terms of impairment in attention or alertness and suggestions were made about future research relating cerebral events and CPT performance.

Received January 13, 1956.

References

1. Hebb, D. O. *The organization of behavior.* New York: Wiley, 1949.
2. Levin, S., Greenblatt, M., Healey, M. H., & Solomon, H. C. Electroencephalographic and clinical effect of prefrontal lobotomy, with consideration of postlobotomy convulsive seizures. In M. Greenblatt, R. Arnot, & H. C. Solomon (Eds.), *Studies in lobotomy.* New York: Grune & Stratton, 1950. Pp. 400–427.
3. Lindsley, D. B. Electroencephalography. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: Ronald. 1944. Pp. 1033–1103.
4. Mettler, F. A. (Ed.) *Psychosurgical problems.* New York: Blakiston, 1952.
5. Mettler, F. A. (Ed.) *Selective partial ablation of the frontal cortex.* New York: Hoeber, 1949.
6. Partridge, M. *Pre-frontal leucotomy.* Springfield, Ill.: Charles C Thomas, 1950.
7. Petrie, A. *Personality and the frontal lobes.* New York: Blakiston, 1952.
8. Schwab, R. S. *Electroencephalography in clinical practice.* Philadelphia: Saunders, 1951.

*Measures of Cognitive Functioning Appropriate for Use in Pediatric Psychopharmacological Research Studies

James M. Swanson, Ph.D.[1]

Introduction

Six types of cognitive tasks, each of which has been used previously in several studies and by several investigators, were selected for description in this paper:

1. Continuous Performance Tasks
2. Choice Reaction Time Tasks
3. Single-Trial Recall Tasks
4. Matching-to-Sample Tasks
5. Paired-Associate Learning Tasks
6. Analogue Classroom Tests

These cognitive tests may be considered to be ordered on a dimension of complexity or difficulty (from "low-level" to "high-level" cognitive tasks, according to Weiss & Laties, 1962), or on a dimension of effortfulness (from "automatic" to "effortful," according to Hasher & Zacks, 1979). Thus, each measures a different aspect of attention or cognition. Each type is available in a computerized and non-computerized version.

The Six Classes of Tasks

1. The Continuous Performance Task (CPT)

The CPT is a monitoring task in which the subject watches an experimenter-controlled sequence of stimuli (e.g., letters) to detect an infrequently occurring target. A "vigilance decrement" in performance occurs within relatively short (5 to 30 min) periods of time. Several versions of the CPT are outlined in Table 2. Detailed descriptions of these CPTs are provided by Nuechterlein (1983). This test has been used to document attentional deficits in hyperactive children (e.g., Sykes et al., 1971, 1973) and children at risk for schizophrenia (Nuechterlein, 1983; Rutschmann et al., 1977). The CPT has a long and distinguished history of use with children, dating back to Rosvold et al. (1956); auditory versions, requiring little equipment (e.g., a tape recorder), to visual versions requiring special purpose projectors and computers are available; it is sensitive to medication effects. However, most versions present a "low-level" task, and the more sophisticated versions may be required to measure psychopharmacological effects on effortful cognitive processes (Nuechterlein, 1983). A related perceptual test, the Span of Apprehension test, is described by Asarnow (1984).

2. Choice Reaction Time Tasks

Choice reaction time tasks have been used to study stages of human information processing (Callaway, 1983; Sanders, 1983; Sternberg, 1969; Welford, 1960). The Sternberg (1966) and Shiffrin and Schneider (1977) scanning tasks, which address four stages of information processing (stimulus encoding, memory comparison, response selection, and response execution), use digits or letters as stimuli. Memory load and display load are the major independent variables, and the dependent measures are reaction time (RT) and error rate. The slope and intercept of a linear RT-load function are interpreted and used to estimate the time required by component stages (e.g., memory comparison time). Sprague and Sleator (1977) used a version of this task with large display loads (up to 15 items), and found that the optimal dose for performance (RT and errors) on it were lower than for subjective rating of behavior. Sergeant and Scholten (1983) and Sergeant (in press) have shown clear differences between normal and ADD/H children in the intercept of the RT-memory load function, but no slope differences, and Callaway

[1]Child Development Center, University of California, Irvine, CA 92717.

*PLEASE NOTE: All references in this paper are presented in a single reference section to be found on pages 1111-1124.

Table 1

Cognitive Tasks Recommended for Use in Pediatric Psychopharmacological Studies

Tests	Source	Selected References
Continuous Performance Tests	Dr. Keith Nuechterlein UCLA-Neuropsychiatric Institute Box 18 760 Westwood Plaza Los Angeles, CA 90024	Nuechterlein, 1983 Sykes et al., 1971, 1973 Rosvold et al., 1956
Choice Reaction Time Tests	Dr. Joe Sergeant Laboratory for Experimental Clinical Psychology Turfsingel 46 Groningen, The Netherlands	Sergeant & Scholten, 1983 Sprague & Sleator, 1977 Sternberg, 1969
Single-Trial Recall Tests	Dr. Herbert Weingartner Lab. Psych. & Psychopath. National Institute of Mental Health Bldg. 31, Room 4C-35 9000 Rockville Pike Bethesda, MD 20205	Weingartner et al., 1980 Rapoport et al., 1978 Craik & Lockhart, 1972
Matching-to-Sample Tests	Dr. Robert Sprague Institute Child Development & Behavior University of Illinois 51 Gerty Drive Champaign IL 61820	Sprague, 1984 Campbell et al., 1971 Kagan et al., 1964
Paired Associate Learning Tests	Dr. James Swanson Child Development Center University of California 19262 Jamboree Road Irvine, CA 92717	Swanson et al., 1983 Gittelman-Klein & Klein, 1975 Conners et al., 1964
Analogue Classroom Tests	Dr. William Pelham Dept. of Psychology Florida State University Tallahassee, FL 32306	Pelham et al., 1985 Stephens et al., 1984 Douglas et al., in press

(1983) and Klorman et al. (in press) have used evoked potential latencies as well as RT to show that methylphenidate does not affect the slope of the RT-memory load function. These findings suggest that ADD/H children have functional, not structural deficits associated with the memory comparison process, and that methylphenidate affects post-evaluative processes, perhaps at the motor selection and execution stages of processing.

3. Single-Trial Recall Tasks

Cognitive psychologists (e.g., Craik & Lockhart, 1972) have developed tests to study specific input (encoding) and output (retrieval) processes of human memory. Lists of words are used as stimuli; the words are typically presented one at a time to the subject, who then attempts to recall them in any order. Independent variables related to encoding and retrieval are established by requiring subjects to process the words in different ways during encoding (e.g., semantically or acoustically), and by guiding the subject with cues (in both encoding and retrieval phases, in one phase but not the other, or in neither phase). Both recall and recognition tests are given, to probe for the availability (recognition) as well as the accessibility (recall) of information in memory (Craik & Tulving, 1975). The studies of Rapoport et al. (1978) and Weingartner et al. (1980) demonstrate the use of this test for assessing the encoding and retrieval strategies of hyperactive children, and the effect of medication on recall. These tests focus on the processing of episodic information in short-term or working memory (one-time recall), and typically do not require "learning" or reorganization of information in semantic memory.

Table 2

Computerized and Non-Computerized Tests Recommended for Use in Pediatric Psychopharmacological Studies

Tests	Computer	Investigator
Computerized		
CPT-Degraded Stimulus	IBM PC	Nuechterlein, Lindgren
CPT-Playing Card	Commodore	Taylor, Lindgren
CPT-AX	Apple IIe	Garfinkel
Span of Apprehension	Apple IIe	Asarnow
Recall	Atari, Apple	Sandman, Swanson
Memory Scanning	Apple II+e,c	Swanson
Pursuit Tracking	Apple II+e,c	Swanson
Matching-to-Sample	Apple IIe	Sprague
PAL Picture-Letter	Apple IIe	Deutsch
PAL Word-Word	Apple II+e,c	Swanson
Arithmetic Tests	Apple II+e,c	Swanson
Non-Computerized	Equipment	Reference
CPT-Auditory	Tape recorder	Campbell et al., 1971
Etch-a-Sketch Tracking	Etch-a-Sketch	Humphries, 1979
Recall Free-Cued	Word lists	Swanson, 1985
Recall Categories	Word lists	Kagan, 1966
Recall Encoding	Word lists	Weingartner et al., 1980
Display Scanning	Letter lists	Neisser, 1967
MFF-20	Notebook	Douglas, in press
PAL Picture-Letter	Slides	Swanson, 1983
PAL Word-Word	Word lists	Douglas, in press
Analogue Spelling	Word lists	Pelham et al., 1985

4. Matching-to-Sample Tasks

The Matching Familiar Figures (MFF) test (Kagan, 1966; Kagan et al., 1964) is a simple version of this test in which a standard or target picture is presented, along with a set of similar pictures which contains one that is identical. The subject's task is to find the exact match. The usual dependent measure is the time-to-first response (a measure of reflection time) and the number of errors made before a match is found. In some studies (e.g., Campbell et al., 1971) stimulant medication has produced a pattern of more reflective responding (longer latencies and fewer errors). In other studies of medication (imipramine and methylphenidate), latency has increased, but errors have not decreased (e.g., Rapoport et al., 1974). Flintoff et al. (1982) used computer-generated forms and recorded eye movements in a Kagan-like MFF test. They found that stimulant medication produced an increase in systematic comparisons of the variants to the standard (an increased selective attention), but no improvement in performance as measured by latency to first responses or errors. Sprague (1984) has developed a sophisticated, computerized version of the matching-to-sample task designed to measure individual differences in strategies as well as levels of performance. It utilizes letter-like stimuli (Gibson figures) which can be systematically transformed to vary difficulty, and computerized monitoring of the subject's selection of stimuli to inspect. Sprague (1984) has shown that the inspection ratio (the number of times each stimulus window is opened per trial) is sensitive to the beneficial effects of low doses, and shows the behavioral toxic effects of high doses in the same way as his earlier memory scanning test (Sprague & Sleator, 1977).

5. Paired-Associate Learning Tasks

In the paired-associate learning (PAL) test, stimulus-response pairs of common, well-known items (e.g., picture-letter pairs, or word-word pairs) are used. The subject's task is to learn the arbitrarily assigned response to each stimulus item, through rote repetition and rehearsal. Each trial consists of a new study-test-feedback sequence, and error scores are calculated for each trial.

Typically, subjects are required to continue the task until reaching a criterion of perfect performance. In one of the earliest series of studies of the effects of stimulant drugs on hyperactive children (Conners, 1973; Conners et al., 1964), and in one of the most recent series of studies (Douglas et al., in press), the PAL test was found to be a good test for distinguishing hyperactive from normal children and for monitoring the effects of stimulant medication. The PAL test has been used by Swanson and Kinsbourne (1976; Swanson, in press-b; Swanson et al., 1978) in more than a dozen studies, and by at least 10 other investigators in studies of ADD/H children (*see* Swanson, in press-b). The paired-associate learning test is a "high level" or "effortful" test, and because of task-specific effects of stimulant medication, lower doses may produce indications of "behavioral toxicity" on this, like on other effortful tests (*see* Sprague, 1984), than on "low level" or "automatic" tests (*see* Rapport et al., in press; Swanson, in press-a). It is one of the best tests for showing differences between ADD/H and normal subjects (Douglas, in press). Medication-induced changes in performance on it are correlated with behavioral changes reported by teachers and psychiatrists (Gittelman-Klein & Klein, 1975). However, some ADD/H children who show clear behavioral improvement may show no improvement or impairment on the PAL test (Swanson et al., 1983).

6. Analogue Classroom Tests

Recently, tests of reading, spelling, and arithmetic have been used to document the effects of medication on academic productivity in the classroom. Pelham et al. (in press; Stephens et al., 1984) used an analogue spelling test, in addition to reading and math tests, to monitor academic production and medication effects in ADD/H subjects. Douglas et al. (in press) have used arithmetic tests to show short-term effects of medication on ADD/H children in the classroom. These tests border on being achievement and psychometric tests, and are for that reason somewhat out of the realm of this paper on cognitive tests for use in pediatric psychopharmacology. However, the analogue spelling test is a laboratory-like classroom test (Pelham et al., in press)

Summary

Table 1 summarizes the six types of cognitive tasks recommended for use in pediatric psychopharmacological studies. These tests are designed to measure different aspects of attention and cognition, but all have been shown to be sensitive to the effects of medications in children, and useful in documenting deficits in clinical groups of children. In general, the "low level" or automatic tests are more sensitive than the "high level" or effortful tests for documenting effects of medication, but the "high level" or effortful tests are more specific than the "low level" or automatic tests for distinguishing groups or subtypes.

The studies by Sprague and Sleator (1977) and Rapoport et al. (1978) are recommended as two of the most important studies in pediatric psychopharmacology of the past decade. The articles by Conners (1973), Gittelman-Klein and Klein (1975), Callaway (1983), and Douglas et al. (in press) are highly recommended and should be useful to anyone planning a study of cognition in pediatric psychopharmacology. Computerized batteries of tests are available from Keith Nuechterlein, (U.C.L.A.) Scott Lindgren (Univ. of Iowa), Barry Garfinkel (Univ. of Minnesota), and James Swanson (U.C., Irvine).

*The Computerized Continuous Performance Test

C. Keith Conners, Ph.D.[1]

The CPT has been used in many studies as one operational definition of "attentional" dysfunction in children. The test consists of a series of visual or auditory stimuli, some of which are designated as targets to which a button-press is required. Usually the test includes an easy version in which a single target is designated, and a more difficult version in which the same stimulus can be a target or nontarget depending upon whether it is preceded by another particular stimulus (such as an "X" when preceded by a "B"). What the test actually measures has never been adequately studied, and investigators are cautioned that such concepts as "sustained attention" and "attention span" are inappropriate for the task when used over brief periods. Generally, the CPT records errors of omission and commission, but not decline in vigilance over time (though there is no obstacle to extending the task and recording such changes).

Unfortunately, no standard version exists. Parameters which need to be considered are total number of stimuli, total test duration, interstimulus interval (ISI), probability of target occurrence, stimulus duration, and method for calculating scores. Omissions and false alarms can be used to calculate signal detection parameters of d′ and Beta, the former measuring threshold sensitivity, the latter the response bias or criterion for making a response. Variations in response style, such as cautiousness or impulsivity, affect the speed-accuracy trade-off. While mechanical or computerized methods are desirable, paper-and-pencil methods such as cancellation tasks or "checking" have been standardized (Keogh & Margolis, 1976).

The following are computerized methods which are readily adaptable to a wide age range for drug studies and sample selection.

1. NIMH Version

This early version test has been developed and standardized upon a normative sample. It presents 800 visual stimuli (3/8" high digits) sequentially presented for 100 milliseconds (ms) on a LED display. The task requires the detection of infrequent two-digit sequences (6 followed by 4). The ISI is automatically reduced by 5% following each correct identification and by increasing the ISI by 5% following each error, thus adjusting the rate according to the subject's performance. d′ and Beta are calculated (Sostek et al., 1980). The apparatus was designed and built by NIMH, but is not available for sale.

An AX version of the task, currently in use by Rapoport and associates at the NIMH, has a sequence of single letters presented on an LED display. The child pressed a button if X appeared, only if preceded by A. Failure to do so was an omission error; pressing the button for a letter outside the critical sequence was a commission error. Children were instructed to work to maximize the stimulus-presentation rate. A total of 800 digits was presented; there were 32 A-X combinations (4%), 96 A (12%), and 96 X (12%). The mean interstimulus interval for the last 400 stimuli formed a third summary score. The child is usually tested for approximately 10 minutes.

The adaptive version of the task in which the speed is increased or decreased by 5% following a correct or incorrect response, respectively, permits testing over a wide age range using the "sure" task. For younger children (6 and below) the adaptive version often becomes too slow, becoming instead an (extremely) de-

[1]Children's Hospital National Medical Center, Washington, DC 20010.

*PLEASE NOTE: All references in this paper are presented in a single reference section to be found on pages 1111-1124.

layed reaction time task. For such subjects, the fixed interval version must be used. While not currently for sale, the NIMH Adaptive AX task is being developed by Michael Gordon and will be available to the public.

2. Klee and Garfinkel (1983)

This version is programmed in PASCAL for the Apple-II computer. Ten letters flash on a video monitor for 130 ms with an ISI of 600 ms. The target is the letter T preceded by the letter S. Ten percent of 500 letters are targets. Errors of omission and commission are calculated.

3. Behavioral Medicine—CPT

This is a program written for the Apple-II computer in BASIC. The stimuli are large alphabetic block letters with a display time of 50 ms. ISI is 1.5 sec. Three hundred trials are presented of which 50 are targets for each task. A training session is optionally usable for standardized practice and instructions. Included with the same diskette is a computerized version of the Sternberg test of short-term memory, and a visual sequential memory test. Available from Behavioral Medicine, Inc., 3904 Cleveland St., Kensington, MD 20895 ($150).

4. Life Sciences Associates Tasks

This is a signal detection task which calculates standard signal detection parameters. The program is written in BASIC and is available for both Apple-II and IBM PC computers. Stimuli are small letters which may be displayed anywhere on the video monitor. Targets are the letter X which occurs as specifiable frequency, duration, and rate amidst distractors (letter Z). The ISI can be varied across a wide range. Available for $49 from Life Sciences Associates, 1 Fenimore Rd., Bayport, NY 11705.

5. The Gordon Diagnostic System (GDS)

This system provides researchers and practitioners with a means of obtaining accurate data on a child's ability to delay and to sustain attention. The portability and ease-of-operation of the GDS has enabled large-scale data collection (Gordon, 1979; McClure & Gordon, 1984).

The GDS allows for the administration of two tasks, the Delay Task and the Vigilance Task, a version of the Continuous Performance Task (Rosvold et al., 1956). A new procedure, the Distractibility Task, is currently being developed to assess the effects of distraction on the child's ability to sustain attention.

A microcomputer software package is available that automatically receives data from the GDS, tabulates the various scores, stores the information, and prints out an interpretive summary of the performance. Included in the report are educational recommendations and suggestions for pharmacotherapy.

Readers interested in further information on the GDS may contact Michael Gordon, Ph.D., at Upstate Medical Center, 750 East Adams Street, Syracuse NY 13210 (Tel. 315-473-8145). Information about purchase of the GDS may be obtained from Clinical Diagnostics, Inc., 300 E. Mineral Ave., Suite 3, Littleton, CO 80122 (Tel. 800-521-4503; 303-795-0438 in Colorado).

All these measure are sold for clinical use.

*Measuring Activity Level in Children

C. Keith Conners, Ph.D.,[1] and Sandra Kronsberg, Ph.D.[1]

"Activity" has become a central concept, both in the understanding of normal development, and of childhood psychopathology. Activity level at a given age, rate of development across the age span, and patterning of activity are parameters which show important relationships with measures of personality, behavioral style, learning, and maladaptive functioning. Activity in man, as in most mammalian species, is a recognizable trait with strong evidence of high heritability (Buss & Plomin, 1975; Buss et al., 1973; Guilford & Zimmerman, 1956; Owen & Sines, 1970; Scarr, 1966; Schoenfeldt, 1968; Thurstone, 1951; Willerman, 1973). It is to be expected that biological variation alone would ensure that some children would be strongly characterized as either under- or overactive, traits which may be predominant in their behavior over much of their life.

However, stable individual differences in activity level are only one part of a more complex picture. Activity level changes over the age span. Almost all methods of measuring activity show a decline of activity level with age, whether by global impression of parents or caretakers, discrete samples of behavior, or mechanical transducers of physical energy. Again, by general principles of behavioral ontogenesis, one can assume that there will be pathologies of development, such as a failure to modulate the level of activity in accordance with normative standards at a particular age. As with pathologies of activity *level*, developmental deviations in activity can simply be the extremes of a normal distribution (in this case, of *rate of change* in activity); or these deviations can be the consequence of a variety of biological insults and environmental stressors.

Activity is not merely a genetic expression of a stable trait, nor simply a parameter of normal development. It is, in fact, not a unidimensional construct, but one which requires specification of environmental variables which act to inhibit, exacerbate, or modulate the behavioral acts of the child—acts which are summed together by one method or another—and labelled "activity." Activity is a resultant vector whose origin may involve the social, educational, and perhaps nutritional environments interacting with the biological substrate of activity. There is great variability in the degree and rate at which raw energy becomes tamed by the socialization process.

It is the environment, most of all, which *patterns* activity level. We may suppose, then, that there will be meaningful categories of environmentally caused pathologies of activity level, which will be best described in terms of an abnormality of *pattern*; of fit between environmental press and temperamentally based activity. Thus, undersocialization is one environmental hazard which may lead to impulse disorders whose expression can resemble hypermotility. Such disorders are, of course, fundamentally different from impulse disorders which derive from the *inability to apply restraint* in the context of normally adequate socialization. Another type of "hyperactivity" is the profound curiosity which sometimes drives the talented child to continually challenge the bounds of boredom, perhaps along the way incurring the displeasure of some adult who finds him "hyperactive." Nor is the underactivity of the dreamy genius the same as the hypoactivity of the forlorn child. These are distinctions which can be made only as part of a pattern of findings, and for this task the human observer far surpasses mere quantitative recordings of changes in physical energy by means of the motion actometer or counts of

[1] Children's Hospital National Medical Center, Washington, DC 20010.

*PLEASE NOTE: All references in this paper are presented in a single reference section to be found on pages 1111-1124.

small units of behavior. Some authors speak of this as *motivated activity*, and relate it to the goals and purposes of the actors within a social context.

It is not surprising that there should be disagreements among different operations for assessing activity level. The human observer may think he or she is as sensitive as the pendulum-driven watch spring which records changes in acceleration. The observer is really sensitive to different dimensions; to the meanings, goals, proprieties, and social codes which constrain behavior. The observer can detect when an event is a reward, a punishment, or a coercion, and grasp the nature of physical or social pain which drives and constrains particular behaviors. Whether the observer is the actor or the other, the patterned appropriateness of activity is best judged by them. When the mother says her child is overactive, she may mean, "For me, in this house, at this time of my life, with my level of patience, and my expectations of behavior, this kid is driving me nuts." Excellent, then, as a measure of social appropriateness and expectation, such judgments may be meaningless with regard to quantitative measures of energy expenditure.

These measures make *independent* contributions to the prediction of behavior. It is perfectly appropriate to ask parents whether their child is up and down at the dinner table a lot, but it is cultural norms which prescribe when a certain level of ambulation during dinner is too much, and which also govern the sanctions that shape the expected behavior. The absolute level of energy consumption is only dimly related to many of these social variables. We may well expect to find, then, that correlations among measures of socially defined *inappropriate* activity and counting methods ("meter reading") of physical energy are low. Barkley and Ullman (1975) found that activity measures tended to correlate with each other, but the relationships were small in magnitude, and differed according to the type of sample. In general, though, hyperactives tended to show a global hyperactivity which generalized across settings, but this was not true for nonhyperactive clinic patients and controls.

We believe, then, that there is a role for quantitative measurement of activity level as a trait whose interactions with drugs is quite meaningful, both as independent and dependent variable. There is also a role for self or other judgments of risk-taking, drive, sensation seeking, and curiosity behaviors which need to be considered equally carefully. Lack of restraint, or inhibitory capability may not be the same thing as innately high drive level towards activity. The rubric of "activity" subsumes many distinguishable notions.

Researchers in psychopharmacology need to be aware of the properties of the various measurement methods and make a judicious choice based upon the requirements of the specific problem and setting. The study of activity is, unfortunately, a methodologic minefield. (Though slightly outdated, one of the best reviews of methodologic issues in activity level research is the comprehensive chapter by Cromwell et al., 1963.)

Rating and Self-Report

Ratings are the method of choice when cost and feasibility are severely limiting factors, and when what is required is an estimate of the fit between the temperamental trait of activity level and the family. Little will be said here regarding these methods since they are discussed elsewhere in this volume.

The Werry-Weiss-Peters checklist (WWP) has been successfully employed as a drug-sensitive measure in a number of studies (Werry, 1968; Werry et al., 1966). This scale records the impressions of parents regarding a child's activity in a number of settings, such as mealtime, television watching, play, sleep, and social settings. The scale has been factor analyzed (Routh et al., 1974), and there is little agreement between activity levels displayed in one home situation and those in another. Rated activity was found to show an almost perfect linear decrease in scores across the age range from 3 to 9 in normal children. Interparent agreement ranged from 0.16 to 0.58 (median = 0.33), and sex was nonsignificant as a variable. Although a similar monotonic decrease in activity was found for direct observation in a playroom, there was no significant relationship between the two types of measures. A similar finding was reported by Barkley and Ullman (1975). These findings imply that one should use this scale not in place of other methods, but as a complement to them. In the absence of more sophisticated (and costly) techniques, it provides a straightforward assessment of activity in the home setting which makes it attractive in psychopharmacologic studies.

The Conners teacher and parent rating scales each have a "hyperactivity" factor which has proven drug-sensitive and diagnostically useful. Along with the 10-item abbreviated scale these methods have been extensively studied, and their ease of use makes them irresistible as "quickie" measures which allow comparison with a large normative database and many previous drug studies. But extensive use has also documented the variable and inconsistent relationship of such ratings to other methods, and caution is recommended in exclusively relying upon ratings since they are sensitive to context, informant, practice effects, age, and sex.

Buss et al. (1980) have provided a Q-sort method for descriptions of children at the preschool level, including energy, restlessness, fidgetiness, caution, aggressiveness, and assertiveness, which shows good validity with actometer measures.

The concept of "impulsiveness" has a close relationship to activity level. Eysenck et al. (1984) have recently published a comprehensive normative study on 1,505 7- to 15-year-old normal children, using a self-report questionnaire. Factor analysis showed that the same factors emerged for boys and girls, and confirmed the expectation that impulsiveness consists of two distinct factors: impulsiveness in action or speech, and venturesomeness (including risk-taking and sensation-seeking). Reliabilities and factor homogeneity are excellent, and the scale may prove useful as a descriptive measure for subject samples, though it seems unlikely to be drug-sensitive given its trait-oriented nature.

Direct Observation

Recording of free-field activity in playroom settings has provided valuable information regarding both normal and hyperactive children (Hutt & Hutt, 1970; Milich, 1984; Pope, 1970; Routh et al., 1974). Generally, these methods require that the room be marked off into grids from which crossings can be calculated. An excellent guide to the development of such methods may be found in Hutt & Hutt (1970).

As described elsewhere in this volume, the observational scale by Abikoff et al. (1980) is recommended as a reliable and valid measure of gross and minor motor behavior within a classroom setting when trained observers are available.

Mechanical Methods

Schulman & Reisman (1959) were among the first to apply the actometer to clinical child studies. The actometer is a wristwatch which has been modified so as to be sensitive to changes in acceleration. The validity of actometer measurement has been seriously questioned by Johnson (1971) who reported that when two actometers were attached to the wrist of someone using a hammer, the more distal actometer had significantly higher readings, suggesting that arm length may be a confounding variable. However, methodological studies have shown that arm length in young children shows no relation to results, and that although recordings from a single actometer are somewhat unreliable, reliability increases rapidly as more than one actometer is measured and several samples are recorded (Eaton, 1983).

Considerable data support the validity of actometer measurement, though there are some conflicting data. Following a suggestion by Bell (1968), Halverson and Waldrop (1973) studied indoor and outdoor play activity in preschoolers, finding significant correlations across settings as well as significant relationships with teachers' ratings of activity. Importantly, the sex differences in activity were detected with the actometer but not with teacher ratings. Sex and age effects were found for staff ratings as well as actometer scores. Actometer and teacher scores were also significantly associated in a study by Buss et al. (1980), who found that actometer measures remained consistent across a 4-year age span. Stevens et al. (1978) also found that actometer scores correlated strongly with mother and trained clinical staff ratings of activity level. Milich (1984) found good intra-subject stability coefficients for actometers over a 2-year period, but only for restricted settings ($r=0.47$) and not free-play.

Kendall and Brophy (1981) examined the interrelationship of teachers' ratings (Conners Hyperactivity factor), a stabilimeter chair, and wrist actometers. The rating factor was significantly but modestly associated with the actometer ($r=0.26$) but not with stabilimeter or behavioral observations. The actometer correlated well with the stabilimeter ($r=0.65$) and the behavior observations ($r=0.53$). The stabilimeter also related well to the behavioral observations ($r=0.58$). It should be noted that the sample was normal school age children ($N=49$), of whom only 10 would have qualified as hyperactive on the Conners scale. Again, age and sex effects upon activity level were noted.

Actometers have been used successfully as a drug response measure by Rapoport et al. (1980), who measured truncal activity in a cognitive testing situation with amphetamine-treated normal and hyperactive boys, and normal adult men. This actometer uses a solid-state memory and allows continuous measurement over time, with readouts at selectable time intervals (Colburn et al., 1976). The sensitivity of this instrument was indicated by the fact that the normal adults showed activity decreases with drug, even though they had low levels to begin with.

One of the most important applications of this actometer was by Porrino et al. (1983a), who used the continously recording actometer over a 1-week period with hyperactives and normal controls. Using recordings from consecutive 1-hour periods, they found that the hyperactives were more active in all situations, including sleep. The breadth of sampling made possible by this method effectively eliminates much of the unreliability which occurs from sampling in short intervals in unrepresentative situations. They found that activity during school was the best overall discriminator of the clinical groups. Importantly, this method allowed the investigators to demonstrate that activity and attentional (CPT) measures were independently contributing to the discrimination between groups, misclassifying different subjects: the motor measure misclassified older hyperactives who tended to move less, while the CPT misclassified younger controls who had difficulty with the CPT. In addition, the measure was highly sensitive to stimulant drug effects, with powerful interaction between situation and effect (Porrino et al., 1983b).

Resources for Actometers

Several sources are available for continuously recording actometers:

1) The NIMH actometers developed by Colburn et al. (1976). These are apparently not directly available but could conceivably be manufactured according to the specifications reported in the original article.
2) Mr. Gary Mathews
   KFM Corporation, Inc.
   c/o Western Psychiatric Institute and Clinic
   3811 O'Hara Street
   Pittsburgh, PA 15232
   (412 624-2353)

Have the rights to the NIH patent for non-government use.

3) Precision Control Design, Inc.
   646 A. Anchors Street
   Ft. Walton Beach, FL 32548
   (904 224-1923)

Advertises complete monitor and microcomputer readout support.

4) Ambulatory Monitoring, Inc.
   731 Saw Mill River Road
   Ardsley, NY 10502
   (914 693-9232)
5) Vitalog Corporation
   1058 California Avenue
   Palo Alto, CA 94306
   (415) 365-1100)

In addition to the cost of the actometer, a microcomputer is required for displaying and analyzing the data.

Several versions of the wrist accelerometer have been available in the past from Timex. However, Timex no longer markets these. They are presently available from:

Kaulin & Willis
282 Watertown Road
Middlebury, CT 06762

The model 100 is a wristwatch which is sensitive to motion along the 6 to 12 o'clock vector. They sell at a cost of $130, with reductions in price for larger quantities. A more useful version for children, the Model 101 motion recorder, is identical to the Model 100 except that the recorder hands cannot be set. This renders the recorder tamper proof from unsupervised children (who generally will wish to try and fiddle with the hands).

To those researchers embarking upon the mechanical measurement of activity level, some cautionary notes are indicated. First, most of the recorders need to be independently calibrated, and in the case of the watch varieties, this may need to be done often. Second, equipment breakdown is frequent and somewhat unpredictable. The electronic versions are particularly sensitive to environmental conditions; and showers, fights, and miscellaneous mishaps of childhood render their unsupervised use both costly and frustrating.

[Re]view of Educational Research
Winter 1977, Vol. 47, No. 1, Pp. 293-318

→ See page 310

The Measurement of the Hyperactive Syndrome in Children

Jonathan Sandoval

University of California, Davis

The hyperactive syndrome in school-aged children has become one of today's major mental health problems. Estimates of the prevalence of the syndrome indicate that from 5% to 20% of the elementary school population are affected (Cantwell, Note 1). Evidence is beginning to accrue that links hyperactivity in childhood with later disorders in adolescence and adulthood (Menkes, Rowe, & Menkes, 1967). In school and at home children with hyperactive behavior are extraordinarily difficult to manage and demand an inordinant amount of a parent's and a teacher's time. Although many children may be treated successfully with stimulant medication, this practice has been hotly debated in the press and in professional publications (Grinspoon & Singer, 1973).

Over the last fifteen years an extensive body of studies on the topic of hyperactivity has accumulated, most of which have either described the condition or have evaluated the effectiveness of various treatment regimens for helping children control and change their behavior. Research will undoubtedly continue in this area as new treatments are devised, e.g., diet regulation (Conners, Goyette, Southwick, Lees, & Andrulonis, 1976), and as new medications come onto the market. Other research trends are the prediction of response to medication and the determination of subgroups of hyperactive children on the basis of etiology or other diagnostic information.

· I wish to thank Carl Spring and Dana Sassone for their assistance in reviewing the manuscript and for their generous suggestions.

In much of the research to date, there has been an effort to use objective measures of different aspects of the behavior of hyperactive children, both to explore the nature of the syndrome and to evaluate changes in children which may have come about as a result of treatment interventions. In many instances, investigators have approached the selection of measures from either an exceedingly narrow perspective or have used what may be termed a shotgun approach. Typically, they have either selected one measure such as activity level, ignoring other aspects of the hyperactive condition such as impulsivity or short attention span; or they have administered a large battery of tests with little thought to the dimensions of hyperactivity being assessed. This review is an examination of the instruments that have been used to measure different aspects of the hyperactive syndrome in children. The objective has been to collect and order information on the measurement of the hyperactivity syndrome in children so that a second generation of research may yield more comparable and complete information to guide practice in educational and medical settings.

Some Caveats

Before proceeding, I believe it is useful to point out some of the difficulties encountered in the execution of the review. The greatest problem has been the wide variety of subjects who have been called hyperactive, hyperkinetic, minimally brain dysfunctional, or learning disabled. The diagnosis of a hyperactive condition is most often made on the basis of reports of hyperactive behavior, not on the basis of hard, unequivocal medical or laboratory findings (Sandoval, Lambert, & Yandell, 1976). As a consequence, there is little way to know if the subjects in one study are equivalent to those in another. Most often, unless the focus is specifically on the epileptic or the retarded, subjects are described as having hyperactive behavior yet have no other medical condition which would give rise to the behavior. Within these boundaries the range of severity and types of behavior is immense. In addition, in different studies subjects vary with respect to socioeconomic status, ethnic background, sex, and age, to mention the most obvious dimensions. In this review, I have grouped together studies in which subjects may have been quite different. I have, however, excluded from the review studies on epileptic and retarded children, for the most part, as well as studies on learning disabled children who were not specifically termed hyperactive.

The reader must also recognize that few investigators of hyperactivity have concerned themselves with the measurement instruments they have used. The bulk of the literature concerns

evaluations of medications for use with hyperactive children. Because of this focus, extensive information on reliability and validity is not present in research reports. In many cases the instruments are used because they are a part of a standard assessment battery given to each child admitted to a clinic, and were not especially selected as instruments to study.

Dimensions of the Hyperactivity Syndrome

The hyperactive syndrome consists of the following pattern of symptoms: high activity level, low attention span, impulsivity, aggressive behavior, excitability, and emotional lability (Cantwell, Note 1). In addition many children have learning difficulties, although others do not. These traits are most commonly listed in clinical descriptions of hyperactive children (Lambert, Windmiller, Sandoval, & Moore, 1976) and are the most usual child characteristics assessed in clinical trials of medication with hyperactive children. Activity level is the most straightforward attribute to measure because hyperactivity in children implies excessive and uncontrolled physical movement. The attentiveness-distractibility dimension is seen by many investigators to be the key factor in describing hyperactive children (e.g. Conners, 1972b; Douglas, 1972). Conners, for one, feels that all other symptoms in hyperactive children can be traced back to a deficit in this area. Impulsivity is often differentiated from distractibility as acting without thinking and has been documented by the Montreal group as being a critical aspect of hyperactive children (see Douglas, 1972). Other investigators have identified poor social relationships, excitability, mood swings, and aggressiveness as being key defining traits of hyperactive children (Conners, 1970).

Organization of the Review

Measures of all these attributes have been used in studies of hyperactive children. The measures themselves can be classified into five types: (1) behavioral ratings, (2) observation schedules, (3) direct physical measurements, (4) simple performance tests, and (5) higher-order cognitive tests. In the following pages, I will review the reliability and validity of instruments of each type as they have been established in a variety of research investigations with hyperactive children.

This review does not exhaust the list of all of the measures that have been used to evaluate hyperactive children. Instead, the most commonly used simple procedures have been emphasized in the review, and procedures using elaborate equipment have been deemphasized.

The Validity Problem

The validity of an instrument for assessing an aspect of hyperactivity in children may be established by its ability to distinguish between normal and hyperactive children, by its ability to distinguish between children taking medication and those receiving a placebo, by its correlations with other concurrent measures, and by its internal dimensional structure. In all cases, the instrument in question must have face validity with respect to one or more of the symptoms of the syndrome.

The first validity criterion is obvious. A measure must discriminate between children labeled hyperactive and children who are not. However, to be a thorough test, the group of hyperactive children must be so categorized on the basis of a score on a measure or on the basis of a completely independent procedure that has nothing to do with the measure being validated. In most instances children are considered hyperactive on the basis of the physician's diagnosis. The fact that an instrument discriminates between two particular groups means little until the instrument can be cross-validated with different groups. Another consideration for validity evidence is that the hyperactive children must be off medication at the time of testing. In many of the studies reviewed, it is not clear that this drug-free condition was obtained.

The second and most common method of establishing validity, that of comparing the scores of medicated children with placebo or nonmedicated hyperactive children, is informative but depends on the assumption that a particular medication is effective in changing hyperactive behavior. The danger here is that most medications are evaluated on the basis of change scores on the instruments. The circularity of definitions may be entrapping.

Aside from the physician's diagnosis of hyperactivity, another form of concurrent validity is the extent to which an instrument correlates with other measures of hyperactivity. This is a useful method of establishing the utility of a measurement technique as long as the second measure has established validity. The fact that two measures of behavior correlate does not mean that either measure necessarily has any relation to hyperactivity unless one measure has been independently validated.

The sensitivity to active medication is just one form of construct validity. But construct validity, always difficult to establish, is even more difficult to establish with the hyperactive syndrome. There is little consensus and less unchallenged research about the condition that permits the prediction of relationships. For example, should hyperactive children perform better on an instrument as they grow older? Several investigators (e.g., Mendelson, Johnson, & Stewart, 1971) have questioned the assumption that children "grow out" of the hyperactive syndrome. Most follow-up studies have not supported this generalization; younger children nevertheless, seem to be scored as more hyperactive than older ones. Most investigators have found many more males to be affected with the condition than females. One might expect sex differences in performance on instruments as a result. If medication affects test performance, one might also expect a variation in test performance with drug dosage. If a drug alters behavior at a low dosage, assuming a ceiling has not been reached, it should show more of an effect at a higher level. This is, of course, a simple view of a complex phenomenon—the working of a drug—and the reasoning may not hold. Given the precautions, various kinds of group differences in performance can add to the evidence concerning an instrument's validity.

Factor-analytic methods yield evidence about the validity of a multidimensional instrument or procedure to the extent that the related items or isolated factors are conceptually meaningful or have face validity. In the case of the hyperactivity syndrome, which is generally considered to have multiple causes, the existence of a number of factors in any complex measure may be expected. Factor-analysis as a technique has subjective aspects so it is important to beware of making strong inferences without information about the population tested and the relative independence of the items rated.

Review of Instruments

Ratings

Behavioral ratings by teachers. Behavioral rating scales are relatively simple and straightforward. It is not surprising that several scales to rate hyperactive children have been developed for use by teachers (Bell, Waldrop, & Weller, 1972; Conners, 1969; Davids, 1971; Greenberg, Deem, & McMahon, 1972; Peterson, 1961).

Conners' Behavior Rating Scale (BRS) is by far the most extensively employed teacher rating scale in research use (Conners, 1969). The scale's 39 items may be summed to obtain a global score (Conners & Eisenberg, 1963; Steinberg, Troshinsky, & Steinberg, 1971), may be grouped into three subscales, classroom behavior, group participation, and attitude toward authority (Conners, Eisenberg, & Barcai, 1967; Conners & Rothschild, 1968; Finnerty, Soltys, & Cole, 1971), or groups of items may be used for subscale "factor" scores of I Aggressivity, II Inattentiveness, III Anxiety, IV Hyperactivity, and V Sociability (Kupietz, Bialer, & Winsberg, 1972; Rapoport, Quinn, Bradbard, Riddle, & Brooks, 1974).[1] Sprague and Werry have normed the factor scoring system on 291 Illinois school children (Sprague, Christensen, & Werry, 1974). Conners has also created a short 10-item form of the scale called the Abbreviated Teacher Rating Scale (ATRS) that is summed to a total (Conners, 1972c; Sprague & Sleator, 1973). The ATRS is made up of many of the same items related to factors I, II, and IV on the BRS, plus two items on mood.

In the longer version, the scale has good test-retest reliability (Conners, 1969). But, test-retest is not a particularly good method of estimating reliability for ratings (see Mischel, 1968). Interrater reliability would have been a better indicator of how consistently the scale may be used.

All of the factors except III, Anxiety, have differentiated between normal children and hyperactive or emotionally disturbed children (Kupietz et al., 1972). The total scores of the BRS (Steinberg et al., 1971) and of the ATRS (Conners, 1972c; Sleator, Von Neuman, & Sprague, 1974) are lower for children receiving the Dexedrine, Cylert, and Ritalin than for children receiving a placebo. In addition, Sleator and her colleagues (1974) found a linear relationship between the score and dosage of Ritalin in their subjects.

When investigators have used the subscale system of scoring, the "classroom behavior," and "attitude toward authority," scores of the BRS have been consistently drug sensitive (Conners et al., 1967; Conners & Rothschild, 1968; Finnerty et al., 1971). The "factor scores" also appear to be drug sensitive to Imipramine, Dexedrine, Ritalin, Cylert, and Chlorpromazine. In six studies (Conners, 1969; Conners, Taylor, Meo, Kurtz, & Fournier, 1972; Kupietz et al., 1972; Rapoport, Abramson, Alexander, & Lott, 1971; Rapoport et al., 1974; Waizer, Hoffman, Polizoes, & Englehardt, 1974), factor IV, Hyperactivity, was drug sensitive in all; factor I, Aggressivity, was sensitive in four studies; factor II, Inattentiveness, was sensitive in four; factor V, Sociability, in two; and factor III, Anxiety, in only one (Conners, 1969).

Although the BRS has been used in conjunction with other measures, intercorrelations are seldom reported. Rapoport et al. (1971), however, have reported that the score on factor IV correlates "playroom activity change" in an observational study. One other investigation has been located shedding light on the BRS's concurrent validity. Saxon and Starnes (Note 2) report low, significant correlation with an actometer reading of arm movement, and even lower, nonsignificant correlations with actometer leg movement and a measure of the amount of time the child was in motion in the playroom. In the same study, the BRS was not correlated with the Davids (1971) scale or the Bell scale (Bell et al., 1972) although these latter two scales correlated .74.

1. Most investigators do not use true factor scores but rather group together items which have a high loading on a given factor on a single scale with no attention to weights.

The factor structure of the BRS seems to be consistent with our notions of hyperactivity except that factor IV includes items reflecting both activity level and impulsivity. In many situations it would be preferable to keep ratings of these two attributes separate. In addition, the factor labeled Anxiety (III) is lacking in validity and adds unnecessarily to the length of the scale. Recently the ATRS rather than the BRT has been used in clinical trials, perhaps because it is shorter and excludes many of the Anxiety and Sociability items. The global score from this rating would probably be a sensitive measure for use in drug trials, but it is too undifferentiated for other uses.

The second rating scale that has good reliability and validity data for hyperactive children is Greenberg's Hyperactivity Rating Scale (HRS) (Blunden, Spring, & Greenberg, 1974; Greenberg, et al., 1972; Spring, Blunden, Greenberg, & Yellin, 1977). The earlier version of this measure consisted of 40 items divided into 10 categories of four items each rated on a 4-point Likert scale. The categories are: Restlessness, Impulsiveness, Distractibility, Low Concentration, Low Perseverance, Irritability, Resentfulness, Cheerfulness, Social Participation, and Verbal Expression. These categories are not indicated on the rating form, and the items appear in a random order. The ten categories when factor analyzed yielded three interpretable factors: I, Hyperactivity, composed of items in the first five categories; II, Hostility, composed of Irritability and Resentfulness category items; and III, Sociability, composed of items in the last three categories (Blunden et al., 1974). This factor analysis was performed on a group of normal kindergarten boys, not a hyperactive sample, and perhaps, as a result, like the Conners BRS, impulsivity is not differentiated in the factor structure from activity level or from attentiveness. Aggressiveness (Hostility) and Sociability, however, are defined by items on the scale, but this last factor has poor interrater reliability.

The earlier version of the HRS was revised by Greenberg and Spring (Spring et al., 1977) to eliminate as much as possible items related to the factors of Hostility and Sociability. The new version consists of eleven categories, each containing three items. New items in the categories of Work Fluctuation, Excitability, Poor Coordination, Fatigue, and Rapid Tempo have been added to the scale; the categories Irritability and Resentfulness have been collapsed to Negativism; and Low Concentration, Cheerfulness, Social Participation, and Verbal Expression have been dropped from the scale with three items related to Social Withdrawal taking their place. Items in the categories Restlessness, Impulsiveness, Distractibility, and Low Perseverance, which contributed most to the Hyperactivity factor were, of course, retained. Spring has collected extensive norms for the new version and, after noting ethnic differences, has reported them for 1,140 white children on each category by sex and grade level. The new version, when it was factor analyzed, did yield the sought-after Hyperactivity factor (51% of the variance) and one other factor labeled Extraversion (15% of the variance). The Interteacher reliability coefficients for six teachers in two classrooms for the Hyperactivity factor (computed by adding together the ratings on Restlessness, Distractibility, Work Fluctuation, Impulsivity, and Excitability) were as low as .15 and as high as .71 for different pairs of teachers. Further evidence for the validity of this version of the scale comes from significant sex differences in ratings on the 11 categories as well as ethnic group and grade-level differences on various categories.

Besides validity evidence from the factor structure, Greenberg's HRS has been shown to be drug sensitive (Yellin, Spring, & Greenberg, Note 3) and, in an observation study (Blunden et al., 1972), the category "impulsivity" correlated with observed impulsivity although other categories were not correlated with parallel observational categories. Data are beginning to accrue on this measure, indicating its utility for use in clinical trials; but for other purposes, like the Conners scale, it may be too undifferentiated to use alone in studying hyperactive children. Although the HRS has not been used as widely as the BRS, it should be considered a contender for future research uses.

One other teacher rating scale has been used in clinical trials of stimulant medication and has sometimes been shown to be sensitive to active medication. This scale, the Peterson-Quay Behavior Problem Checklist (Peterson, 1961; Quay & Peterson, 1967), perhaps because it was not developed specifically to study hyperactive children, has two global factors, Conduct Problem and Personality Problem. The 66 items may be totaled or placed in subscales including conduct, personality, and immaturity. Knights and Hinton (1969) did not find the total drug score to be drug sensitive, but others have found the total score to be correlated with the activity recorder (pedometer) reading (Victor, Halverson, Inoff, & Buczkowski, 1973). Other than these pieces of information, there is very little reliability or validity evidence about the scale for hyperactive children.

Behavior ratings by parents. The rating scales that have been developed for use by parents are not promising devices for use in

assessing attributes of hyperactivity over time. They have not been shown to have the same degree of reliability or validity as the teacher rating scales. Conners (1970) (Conners, Rothschild, Eisenberg, Schwartz, & Robinson, 1969) has constructed for parents a lengthy symptoms rating scale that has two versions, one of 73 items and the other of 93 items. Conners found two factors in the scale, hyperkinetic symptoms and neurotic symptoms, and has shown that the first of these is drug sensitive to Dexedrine (Conners et al., 1969). In his 1970 paper in *Child Development*, Conners reports the scale as being made up of six factors, I, Aggressive Conduct Disorder, II, Anxious-inhibited, III, Antisocial, IV, Enuresis, V, Psychosomatics, and VI, Anxious-Immature, but in his 1972 paper (Conners et al., 1972) he claimed eight factors, adding an Impulsiveness factor, an Obsessive factor, and a Hyperactivity factor and dropping the Enuresis factor. All but the Obsessive factor of a short form of the scale were influenced by Dexedrine and Cylert treatment (Conners et al., 1972). Arnold and Smeltzer (1974) have also factor analyzed a version of this scale. The attractiveness of this instrument is clouded by the inconsistent descriptions of its nature. Moreover, its length, its unknown reliability, and the technical nature of the wording (Sprague & Sleator, 1963) make it unwieldy.

The most widely used parent rating scale is the Werry-Weiss-Peters Home Activity Rating Scale (HARS) (Werry, Weiss, Douglas, & Martin, 1966; Werry, 1968). The Home Activity Rating scale contains 31 items in seven settings: (1) During meals, (2) Watching television, (3) Doing homework, (4) Playing, (5) Sleeping, (6) Behaving away from home (except school), and (7) Behaving at school.

The HARS was originally designed to be completed by mental health professionals during an interview with a child's parents. Using this method, two long-time colleagues were able to demonstrate an interrater reliability of .90 (Werry et al., 1966). In many studies, however, the parents themselves have completed the scale (e.g., Knights & Hinton, 1969; Rapoport et al., 1971). Routh, Schroeder, and O'Tuama (1974), using a shortened 22-item questionnaire version, obtained median interparent correlations of .33 for 140 pairs of parents of normal children. This figure is comparable to the low interteacher correlations obtained by Spring (Spring et al., 1977). The HARS measure has been shown to be drug sensitive (Conners et al., 1969; Knights & Hinton, 1969; Rapoport et al., 1971; Werry et al., 1966). According to Knights and Hinton, the items most significantly changed under Ritalin were related to distractibility and activity level.

Routh and his group (Routh et al., 1974) factor analyzed their short form of the scale and obtained seven factors that accounted for 67% of the variance. They labeled the factors Television-watching Behavior, Bedtime and Sleeping Habits, Mealtime Behavior, Attention-Getting Behavior, Hyper Verbal Behavior, General Restlessness, and Behavior at Play. It is interesting to note that parents in this study did not view hyperactivity as uniform over all settings for their children. This finding does not coincide with many notions about the consistency of hyperactive behavior across settings. In a separate study, Routh (Note 4) was also able to demonstrate that the parents of hyperactive children rated their children higher than parents of normal children, but found no relationship between ratings and observed hyperactivity in the playroom.

Some studies (e.g., Greenberg et al., 1972; Knights & Hinton, 1969; Millichap, Aymat, Sturgis, Larson, & Egan, 1968; Winsberg, Bialer, Kupietz, & Tobias, 1972; Yellin et al., Note 3) have had parents complete a teacher-rating scale. When this has been done, the parents are able to detect change when their children have been receiving active medication. It would be informative to learn if the same factor structure defines behavioral dimensions when the parents completed the rating scale as when teachers did so. Although the literature generally reports parent rating measures as being useful, it may be the case that many negative results go unreported.

Sleator and Sprague (1974) reported, "Parent behavior ratings did not vary significantly with dosage level nor did they distinguish between drug and placebo conditions. This was true for both school days and weekend days. Parental insensitivity to drug effects has been replicated in all of the authors' studies" (p. 29). In another article, Sprague and Sleator (1973), argued that parents "may not be a reliable source of information as to the effects of the psychotropic drug on the behavior of their children" (p. 730). The items from parent-rating forms may not reflect home behaviors that would be affected by medication, or it may be that the effects of medication are not visible in the home. Parents do not see their children in a structured, demanding situation where they must attend; also they interact with their children after school when the effects of the medication may have worn off. Other than the studies reported above, there are no other reports of the reliability or validity of parent ratings. If they are to be used, the HARS seems to be the most useful.

Criticisms of Ratings

Although ratings are ubiquitous in studies of hyperactivity, they are subject to substantial criticism. One set of criticisms springs from the way ratings are incorporated into experimental designs. In most studies teachers or parents are asked to rate a child "blind," not knowing whether or not he is in a placebo group or an active medication group. The purpose of concealing the status of the child is to obtain an unbiased rating. But usually the teacher is asked to rate only one child in the class and cannot avoid becoming more sensitive to that child's behavior. As a result, most studies show the "Hawthorne" or "Placebo" effect in the fact that, compared to the initial rating, the placebo group is rated as improved (e.g. Knights & Hinton, 1969). This placebo effect is particularly true of parent ratings (Rapoport et al., 1971). The change in the placebo children may represent a real change attributed to the attention of parents and physician, but it may also result from changes in teacher sensitivity to the rated characteristics. The difference between the placebo group and the active drug group may be statistically significant, but the real difference in terms of meaningful behavior differences may be quite small.

Most of the ratings consist of leading questions (Stewart & Haller, 1975) and cover only symptoms that are clearly negative, none that are positive or positively stated. A better method of collecting rating data would be to ask teachers to rate the entire class or a subgroup of the class, including the hyperactive child, on a scale that contained items related to hyperactivity and items unrelated to hyperactivity.

Ratings may also be improved by providing raters with more concrete descriptors than the usual Likert scale options of "Often," "Somewhat," or "Not at all" (Mischel, 1968). It is not surprising that interrater reliabilities are as low as they are, given the vague terms used in most scales.

Observation Schedules

From time to time investigators have used observational techniques to study hyperactive children, usually to assess the effects of stimulant medication. Observers have most often counted the occurrence of various categories of events. Observations have been made in classrooms (Blunden et al., 1974; Doubros & Daniels, 1966; Patterson, James, Whittier, & Wright, 1965; Sprague, Barnes, & Werry, 1970); and in play settings (Ellis, Witt, Reynolds, & Sprague, 1974; Rapoport et al., 1971; Schleifer, Weiss, Cohen, Elman, Cvejic, & Kruger, 1975; Victor et al., 1973; Whitehead & Clark, 1970) with observers either making frequency counts on the spot or later through viewing a film or videotape (Ellis et al., 1974; Lee & Hutt, 1964; Victor et al., 1973). The observation variables consisted of high inference global ratings by observers in some instances (Rapoport et al., 1971; Victor et al., 1973), but usually the variables were counts of the occurrence of specific low inference behaviors recorded every 5 to 30 seconds.

The rating scales generally are multidimensional and assess several different child attributes. Activity level has been observed by such techniques as dividing a playroom into equal areas, then counting the number of times a child moved across the imaginary grid (Rapoport et al., 1971, 1974; Routh et al., 1974). Other measures of activity level were "Velocity of Movement" (Ellis et al., 1974), "Restlessness" (Blunden et al., 1974), "Number of times up and out of seat" (Schleifer et al., 1975; Sprague et al., 1970), "Whole Body Movement" (Whitehead & Clark, 1970), and "Vigor of Play" (Victor et al., 1973). Attentiveness was measured by counts or ratings of such variables as "Length of Visit to Play Equipment" (Ellis et al., 1974), and "Change of Activity" (Rapoport et al., 1971). Impulsivity has been included only on one observation instrument (Blunden et al., 1974). Mood, excitability, and sociability have been observed by such categories as "Cheerfulness," "Social Participation," "Verbal Expression" (Blunden et al., 1974), and "Withdrawal" (Whitehead & Clark, 1970). Aggression has been measured by counts of "Aggressive Acts" (Schleifer et al.) and "Physical Contact" (Sprague et al., 1970). No single observation instrument has been constructed that measures all hyperactive child attributes, however.

The interrater reliabilities reported have been high (r = .80—1.00) for such low inference behaviors as number of grid crossings, change of activity, aggressive acts, or times child moves away from a work table. Reliabilities were moderately high (71—78% agreement) for moderately inferential judgments such as beginning of "impulsive acts." Highly inferential observations such as "mood" have shown lower interrater agreements (50% agreement or lower).

Two studies attempted to identify hyperactive children from normal children through observations. Schleifer and his colleagues (1975) were unable to observe differences between normal and hyperactive children in a free play setting, but when the preschool children were moved to a structured setting where they were required to play at a table under supervision, the observational measures of activity level and aggressive acts did characterize hyperactive children. Victor and his colleagues (1973) observed elementary school-aged boys in a free play setting but not in a structured situation. Their observation ratings of "hyperactivity" and "vigor of play" did not identify the hyperactive children from controls, although in the same study the activity recorder reading did. Routh et al. (1974), however, were able to establish that grid crossings may be used to differentiate hyperactive from normal children both under instructions to play freely and under instructions to remain in one area.

When observational methods have been used in evaluations of psychoactive drugs, in some studies, observers have detected changes in behavior induced by amphetamine-like drugs, but not tranquilizers (Rapoport et al., 1971; Routh, 1975; Schleifer et al., 1975; Sprague et al., 1970; Whitehead & Clark, 1970). Other studies with Ritalin have not found the observational methods capable of distinguishing between actively medicated and placebo groups (Ellis et al., 1974; Rapoport et al., 1974). The paradox may be resolved in that those observational studies of stimulant-medicated children in structured situations have noted differences, whereas those studies of children in informal situations have not, leading Ellis et al. (1974) and Sroufe (1975), to conclude that stimulants reduce activity level and increase attention capacity in structured situations that demand concentration but do not affect free play behavior.

In some instances, the observational variables have shown good concurrent validity. Activity level observations have correlated with the activity recorder (Victor et al., 1973), attention observations have correlated with teacher ratings of activity level (Rapoport et al., 1971), and impulsivity observations have correlated with teacher ratings of impulsivity, sociability, attention, and activity level (Blunden et al., 1974). In the Blunden study, however, observations of Restlessness, Irritability, Cheerfulness, Social Participation, and Verbal Expression did not correlate with teacher ratings, and in Routh's studies grid crossings did not correlate with parent ratings (Routh et al., 1974; Routh, Note 4).

Observational methods are valuable tools for studying hyperactive children. They are, however, quite costly. Observer training must be extensive to obtain high interrater reliability. Future investigators must be careful to note the context of the observation and should include operational observation categories covering all the attributes of hyperactivity.

Measures of physical activity. Millichap and Fowler (1967) in their methodological review of studies on hyperactivity suggest that direct measures of activity be included in evaluation studies of hyperactive behavior. Direct measures of activity in children have been taken by fitting a room with photoelectric cells (Ellis & Pryer, 1959) or ultrasonic sensors (McFarland, Peacock, & Watson, 1966; Saxon & Starnes, Note 2) by placing a child's desk on a carefully suspended platform (Foshee, 1958) or by placing a radio transmitter in a helmet worn by the child (Davis, Sprague, & Werry, 1969; Herron & Ramsden, 1967). Most of these devices have been used with retarded children, however. Investigators of hyperactive children have commonly used three devices: the actometer, the activity recorder, and the stabilimetric cushion. These are all mechanical devices for sensing movement. They are, respectively, a modified self-winding wristwatch (actometer), a two-dimensional pedometer-type device attached to the child's shirt back (the activity recorder), and a cushion embedded with sensitive microswitches to detect any squirming while seated. The actometer is very reliable when attached to a machine (Saxon & Starnes, Note 2; Schulman, & Reisman, 1959); when attached to a child, however, Johnson (1971) concluded the device was unreliable. Rapoport et al. (1971) noted that the horizontal and vertical measures of the activity recorder correlate .49, which indicates only moderate reliability for this instrument. These measures, as a result, must be regarded as having questionable reliability.

The actometer has been used in three clinical tests of medications by Millichap (Millichap & Boldrey, 1967; Millichap et al., 1968; Millichap & Johnson, 1974) with mixed results. In his earlier study with Boldrey, he found an *increase* in activity in a small number of subjects taking Ritalin compared to subjects taking a placebo. In his second study (Millichap et al., 1968) the instrument did not distinguish between children on medication and the same subjects taking a placebo although scores did seem to distinguish between children on and off medication. In his 1974 study the children most active before treatment with Ritalin were less active on medication, whereas some children (28%) displayed increased activity under Ritalin therapy. In general these subjects who were more active had been less active before treatment. In the 1968 study, however, the most active children at the start did not have different actometer readings under Ritalin.

The activity recorder has been validated with hyperactive children in that it distinguishes between teacher-judged hyperactive and nonhyperactive boys (Victor et al., 1973), but it does not appear to be drug sensitive (Rapoport et al., 1971). The Stabilimetric Cushion not only differentiates between groups (Sykes et al., 1971), but it also seems to be drug sensitive to the stimulant Ritalin although not to the tranquilizer Thioridazine (Sprague et al., 1970; Sprague et al., 1974).

Although these three measures are superficially similar, their use in research settings has been quite different. Both Actometer and Activity Recorder measurements have been made in play situations or over a 24-hour period. The Stabilimetric Cushion, however, has been employed with children engaged in engrossing psychological tests. As Sroufe (1975) has pointed out, it may well be that stimulant medication has the effect of increasing activity in situations demanding little concentration but decreasing activity in settings requiring attention. The results with the three measures might have been similar if they had been used in the same experimental contexts.

Because of the possibility of instrument unreliability, the expense of the equipment, the obtrusive nature of some instruments, and the problem of determining or controlling the context of measurement in a field study, direct measures of physical activity are less attractive than other diagnostic instruments.

Simple Performance Tests

Common visual-motor or performance measures. Four standard perceptual-motor measures have been used in the assessment of hyperactive children: the Bender Visual Motor Gestalt Test, the Frostig Developmental Test of Visual Perception, the Human Figure Drawing Test, and the Porteus Maze Test. Since these measures are often used by psychologists in the schools, they are of particular interest although other performance measures also have been used.

The Bender Gestalt, commonly scored with the Koppitz system, seems to be a measure of both perceptual accuracy and motor coordination. In addition, receiving a low (good) score on the test requires some planning and monitoring of behavior and attention to detail. Of the various aspects of hyperactivity, the Bender score is perhaps loosely related to attention and impulsivity. Most investigators, however, have included it in their test batteries as a measure of motor-coordination. (Others, such as Page, Bernstein, Janicki, & Michelli, 1974, have used the Lincoln-Oseretsky as a measure of motor development.) Large scale studies using the Bender Gestalt have not shown the measure to be sensitive to medication (Conners & Rothschild, 1968; Conners et al., 1969; Conners et al., 1972; Conrad, Dworkin, Shai, & Tobiessen, 1971; Greenberg et al., 1972; Knights & Hinton, 1969; Millichap et al., 1968; Rapoport et al., 1971; Winsberg et al., 1972) or to differentiate normal children from hyperactive children (Palkes & Stewart, 1972). Dykman and his colleagues (Dykman, Ackerman, Peters, & McGrew, 1972), however, were able to demonstrate that the test could differentiate between learning disabled children and normal children but warned that the difference occurs most in younger children who draw the figures "too rapidly, leading to careless reproduction (p. 47)." Evidently the test has value to the extent that it measures impulsivity in some children; however, there are perhaps better measures of impulsivity.

Another test of visual-motor perception and coordination is the Marianne Frostig Developmental Test of Visual Perception. The Frostig yields an overall perceptual quotient based on scores on five subtests: eye-motor coordination, figure-ground discrimination, form constancy, position in space, and spatial relations. Factorial studies of the test indicate that the items are tapping a single ability, global visual perception skills (Mann, 1972). In three out of four instances where the Frostig was used, the perceptual quotient was drug sensitive (Conners, 1972c; Conners et al., 1969; Conrad et al., 1971). In the fourth study (Millichap et al., 1968), the figure-ground subtests, differentiated between active medication and placebo. This subtest was not drug sensitive in the studies by Conrad and Conners and his associates, but tests, Position in Space (Conrad et al., 1971), and Spatial Relations (both Conrad et al., 1971 and Conners et al., 1969) were drug sensitive. (It should be noted that Conrad selected his subjects on the basis of a low Frostig perceptual quotient.) The Frostig seems to be reflecting some aspects of the attributes of hyperactivity. Inspection of the test indicates that attention to detail in the face of distracting stimuli is at the core of the figure ground test, but doing well on all of the tests depends on refraining from responding impulsively.

The Human Figure Drawing test has long been used as a nonverbal measure of intellectual functioning as well as a clinical tool for assessing various intra- and interpersonal dynamics. The test, when scored by the Harris-Goodenough method, gives a child credit for recollecting and including in his drawing various details of a human figure. When the performance on the drawing is given an IQ score, the test was sometimes drug sensitive (Conners, 1971; Millichap et al., 1968) and sometimes not (Greenberg et al., 1972; Page et al., 1974; Rapoport et al., 1971; Rapoport et al., 1974). When the measure is used in clinical trials with hyperactive children, presumably it is used as a measure of attention to detail rather than as an intelligence test. There is no reason to suspect that short-term drug therapy would affect ability per se (Conners, 1972a), although it may affect attention, concentration, and impulsivity, which would lead to improved performance on the measure. This notion is supported by the findings of Palkes and Stewart (1972) who found that hyperactive children scored lower on the Figure Drawing Test than normal subjects when WISC IQ was partialled out. In addition, in one study (Millichap et al., 1968) the drug effect was strongest for children who had initially scored low on this measure, but were otherwise average in intelligence.

Two investigators have departed from the Harris-Goodenough scoring method with mixed results. Conners (1971) found a "communication organ" score to be drug insensitive. Crowe (1972) using a "formal accuracy" score, found the Figure Drawing test could distinguish hyperactive children from normal children and was drug sensitive. Presumably Crowe's results also reflect attention to detail.

Although the Human Figure Drawing may be useful in drug studies as a measure of attention rather than intelligence, there are more direct measures of attention available for use in following the progress of hyperactive children. The clinical use of the test will probably continue in the search for new understandings about these difficult children.

The other common standardized performance test used to assess hyperactive children's behavior is the Porteus Maze Test. This performance test consists of a series of mazes graded in difficulty that yield an IQ score (variously known as the quantitative score, the Test Quotient, the Test Age score and the T-score), and a Qualitative or Q-score. Both scores have been used to evaluate hyperactive children. Porteus maintains that the IQ score is a measure of foresight, planning, and the ability to profit from experience. A reviewer (Horn, 1972) describes this score as the "capacity or inclination to sustain attention in the face of the difficulty involved in resolving moderately complex spatial relations" (p. 756). The T-score has been shown to differentiate hyperactive from normal subjects (Spring et al., 1976) and in clinical trials with cerebral stimulants, the T-score is drug sensitive (Conners, Eisenberg, & Sharpe, 1963; Conners & Rothschild, 1968; Conners et al., 1969; Conners et al., 1972; Epstein, Lasagna, Conners, & Rodriguez, 1968).

The Q-score, on the other hand, is generally judged to be "indicative of impulsivity or impetuousness—an inability or unwillingness to refrain from behaving in a hasty, slapdash manner" (Horn, 1972, p. 756). A child may receive a high Q-score because of several different types of performance such as going too rapidly, not following instructions, or general carelessness in motor performance. Some of these behaviors fall into the category impulsiveness, others do not. Thus a high Q-score may reflect various qualitative behaviors in contrast to a high T-score, which is purported to be a measure of planning or attention.

When the Q-score has been used as a variable in clinical drug studies, it has not been found to be drug sensitive (Epstein et al., 1968; Rapoport et al., 1974). Palkes, Stewart, and Kahana (1968) in an interesting study illustrated that hyperactive boys given training in verbal mediation to inhibit impulsive behavior showed improvement as reflected in the Q- and T-scores of the test. This was, however, the only evidence of construct validity discoverable.

In sum, these four common psychological tests seem only indirectly to measure attention and/or impulsivity. They are not, of course, measures of aggression, sociability, activity level, excitability, or mood.

Laboratory measures. More specialized performance measures have been explored in laboratory settings. A group of investigators at Montreal, for example, has looked at measures of automatic functioning that are subject to distraction; the Stroop Color Distraction Test and the Santostefano and Paley Colour Distraction Test. These tests require the subject to read a color-name in the presence of a distracting factor, the color of the ink. Measures from these tests of susceptibility to interference and distractibility did not differentiate hyperactive children from normal controls (Campbell, Douglas, & Morgenstern, 1971; Cohen, Weiss, & Minde, 1972) and were not drug sensitive (Campbell et al., 1971; Waizer et al., 1974). In the Campbell study, however, the hyperactive subjects did make more errors of commission and made fewer self-corrections on the Color Distraction test than normal children did and improved in these respects on drugs. Perhaps this later index of what may be impulsivity from the tests has more utility than the indices of attention-distractibility. It may be that hyperactive children have short attention spans but are no more subject to interference *while they are attending* than normal children (Douglas, 1972).

Another set of laboratory performance measures used by a number of investigators center around speed, reaction time, and errors on simple tasks which demand the child's attention for various lengths of time (Campbell et al., 1971; Cohen, Douglas, & Morgenstern, 1971; Conners & Rothschild, 1968; Jacobs, 1972; Spring, Greenberg, Scott, & Hopwood, 1973; Sykes, Douglas, & Morgenstern, 1972, 1973; Sykes, Douglas, Weiss, & Minde, 1971). Anderson (Note 5) for example, has adapted a vigilance task similar to one used in human factors research, where a subject views a pair of colored lights flashing every two seconds and responds only when a certain combination appears. These tasks seem to be reliable (Test-retest r = .50—.87, Sykes et al., 1973), but fatigue and practice-effect must be taken into account when they are used. Most of these measures differentiate between normal and hyperactive children, particularly on tasks which demand complex decisions over an extended period of time. In these situations, hyperactive subjects do not differ in speed of response but make somewhat fewer correct responses and make significantly more errors of a type which results from impulsive responding (Sykes et al., 1971, 1972, 1973). Medication seems to improve functioning on these tasks when actively medicated children are compared with placebo children. Sroufe (1975) points out that the drug effects can be explained by a quicker deterioration in the placebo group's performance over time. As a result these instruments may be considered a measure of sustained attention.

Most of these techniques use sophisticated or elaborate equipment not suited to field studies. They are, however, important

measures of attention. Spring, Yellin, and Greenberg (1976) have suggested two alternative measures on which hyperactive children differ from normal children, and which could be used in field studies. They are extended versions of the Digit Span Forward and Coding subtests of the Wechsler Intelligence Test for Children. As they are longer than the versions on the standard WISC, each becomes a measure of short-term memory and sustained attention, but one, Coding, is drug sensitive, and the other, Digit Span, is not. Spring argues that Digit Span may tap an aspect of attention and memory that is an enduring characteristic of hyperactive children whether or not they are treated with medication.

Higher-Order Cognitive Test Performance

Numerous studies have related hyperactive behavior to cognitive test performance. Although other tests have been used, I will concentrate on the Wechsler Intelligence Scale for Children (WISC), the Matching Familiar Figures Test (MFF) and the Embedded Figures Test (EFT). The WISC, of course, is a widely used measure of intellectual functioning that is routinely administered to children experiencing difficulty in school or entering child guidance clinics. As a result, it has been included in many studies on medication, not as a measure of intelligence, usually, but as an index of improvement in attention and concentration. The MFF and the EFT are measures of the cognitive styles of Reflection-Impulsivity and Field Dependence-Independence.

Because of its common use as a clinical instrument, it is surprising that more investigators have not sought subtest patterns on the WISC to use in differentiating hyperactive children from other populations as well as examining the Performance IQ score, the Verbal IQ test score, and the Full Scale IQ score, for different groups. Keogh and her associates (Keogh, Wetter, McGinty, & Donlon, 1973) divided the WISC subtests into a Verbal-Comprehension score made up of the Information, Vocabulary, and Comprehension subtests; an Analytic-Field-Approach score, made up of Object Assembly, Block Design, and Picture Completion; and an Attentional-Concentration score, made up of Arithmetic, Digit Span, and Coding subtest. Hyperactive subject's scores were depressed on the Attentional-Concentration score, but not on the other scores (cf. Spring et al., 1976).

In another study examining subtest patterns, Palkes and Stewart (1972) found that Similarities, Picture Completion, and Mazes scores of a control group were significantly higher than those of a group of hyperactive children. The IQ scores were also higher.

In clinical trials research, the WISC has been used often but with mixed results. The Full Scale and Verbal IQ measures seem to fluctuate as functions of Performance IQ. Children under medication scored better than children under placebo on the performance scale score in some investigations (Epstein et al., 1968; Knights & Hinton, 1969; Page et al., 1974), but not in others (Conners et al., 1969, Conners et al., 1972; Conrad et al., 1971; Finnerty et al., 1971). When the performance score did improve, Knights and Hinton established that only Picture Completion, Block Design, and Coding had changed, whereas all subtests were higher (no significance tests reported) in the Epstein study. The Coding test here emerged as being drug sensitive in at least one study and was related to an Attentional-Concentration factor in Keogh's division. Here is additional support for using a Coding task as a measure of attention. The use of the WISC itself, however, as a measure of attention and concentration is neither appropriate nor justified, considering the availability of other measures.

The MFF Test was designed by its author Jerome Kagan (Kagan, Rosman, Day, Albert, & Phillips, 1964) to measure the cognitive style impulsivity-reflectivity. An extensive literature exists establishing the construct validity of this instrument. In four studies (Campbell et al., 1971; Cohen et al., 1972; Rapoport et al., 1974; Schleifer, et al., 1975) the MFF has been shown to differentiate hyperactive children of various ages from control children. In addition, in all four studies the impulsivity score has proved to be influenced by active psychostimulants. This is one measure of impulsivity that is consistently useful for studying hyperactive children.

Field independence-dependence, assessed by the EFT (Witkin, 1959), has also been used in research on hyperactivity. The EFT is said to measure one aspect of impulse control, distractibility. Field independent children, the argument goes (Campbell et al., 1971) are better able to act on a problem (locate a figure) in a confused and distracting context (other lines forming a "ground") than are field-dependent children who are presumably distracted. Hence the EFT is used as a measure of attention-distractibility. There is some evidence that hyperactive children are more field dependent than control subjects, but a child's scores on the test does not seem to be influenced by active medication (Campbell et al., 1971; Cohen et al., 1972; Schleifer et al., 1975; Winsberg et al., 1972). As with other instances where attention is assessed in the presence of distractors, there is little drug effect. The EFT is probably not as unequivocal a measure of

attention-distractibility as other tests, but it remains an option for the investigator.

Conclusion

It is evident that most measures used by investigators to study hyperactive behavior are still in a developmental phase. For the majority of the instruments reviewed, reliability coefficients have not been calculated for subjects who are hyperactive. In addition, validity for the measures has been established primarily by only two methods: ability to discriminate between known groups (hyperactive children and normal children) and sensitivity of the instrument to changes in behavior following medication. Considering the nature of the conclusions being made on the basis of data from these measures, they deserve more study.

Campbell and Fiske (1959), in their important article on convergent and divergent validity, point out considerations that should guide research. It is important that different measures be used in evaluation studies and that the correlations and relationships between instruments be examined with the utmost care. Investigators might well consider using teacher and parent ratings, observations, physical measures, simple performance tests, and higher-order cognitive tests in a single study. At the same time they should select measures of activity level, attention-distractibility, impulsivity, aggressiveness, emotional liability, excitability, and sociability from within these measurement techniques. The quality of future research may be improved by the careful planning of a measurement system which employs different kinds of measures of various attributes.

The instruments reviewed also hold promise for clinical practice, both as aids to diagnosis and as tools for evaluation. Selection of a particular instrument by a clinician might be governed by the range of treatment regimens he or she plans to employ. Physicians using medication would obviously attend to the drug sensitivity of an instrument in deciding its appropriateness, whereas therapists using behavior modification techniques would be more interested in instruments that yield accurate baseline information rather than whether the instrument was drug sensitive. The practitioner must be aware of the problems of reliability and the fact that satisfactory norms have been produced for only a handful of the instruments. Limitations aside, the clinician should seek information from sources other than personal observation and an informal interview of parents.

Reference Notes

1. Cantwell, D. P. *The hyperkinetic syndrome.* Chapter to appear in *Recent advances in child psychiatry*, edited by M. Rutter and L. Hersov.

2. Saxon, S. A., & Starnes, K. D. *Hyperactive children: Rating scale validity, parent training, and punishment effects.* Paper presented at the meeting of the American Psychological Association, Chicago, 1975.
3. Yellin, A. M., Spring, C., & Greenberg, L. M. *Effects of imipramine and methylphenidate on behavior of hyperactive children.* Submitted for publication, 1975.
4. Routh, D. K. *Validity of open field locomotion as a measure of hyperactivity.* Paper presented at the meeting of the American Psychological Association, Chicago, 1975.
5. Anderson, R. P. *The vigilance task: A computer based technique for assessing hyperkinesis.* Paper presented at the meeting of the American Psychological Association, Chicago, 1975.

References

Arnold, E., & Smeltzer, D. J. Behavior checklist factor analysis for children and adolescents. *Archives of General Psychiatry,* 1974, *30,* 799-804.

Bell, R., Waldrop, M., & Weller, G. A rating system for the assessment of hyperactive and withdrawn children in preschool samples. *American Journal of Orthopsychiatry,* 1972, *42,* 23-24.

Blunden, D., Spring, C., & Greenberg, L. M. Validation of the classroom behavior inventory. *Journal of Consulting and Clinical Psychology,* 1974, *42,* 84-88.

Campbell, D., & Fiske, D. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin,* 1959, *56,* 81-105.

Campbell, S. B., Douglas, V. I., & Morgenstern, G. Cognitive styles in hyperactive children and the effect of methylphenidate. *Journal of Child Psychology and Psychiatry,* 1971, *12,* 55-67.

Cohen, N., Douglas, V., & Morgenstern, G. The effect of methylphenidate on attentive behavior and autonomic activity in hyperactive children. *Psychopharmacologia,* 1971, *22,* 282-294.

Cohen, N., Weiss, G., & Minde, K. Cognitive styles in adolescents previously diagnosed as hyperactive. *Journal of Child Psychology and Psychiatry,* 1972, *13,* 203-210.

Conners, C. K. A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry,* 1969, *126,* 884-888.

Conners, C. K. Symptom patterns in hyperkinetic, neurotic, and normal children. *Child Development,* 1970, 667-682.

Conners, C. K. The effect of stimulant drugs on human figure drawings in children with minimal brain dysfunction. *Psychopharmacologia,* 1971, *19,* 329-333.

Conners, C. K. Pharmacotherapy of psychopathology in children. In H. C. Quay & J. L. Werry (Eds.), *Psychopathological disorders of childhood.* New York: Wiley, 1972. (a)

Conners, C. K. Stimulant drugs and cortical evoked responses in learning and behavior disorders in children. In W. L. Smith (Ed.), *Drugs, development and cerebral function.* Springfield: Thomas, 1972. (b)

Conners, C. K. Symposium: Behavior modification by drugs: II Psychological effects of stimulant drugs in children with minimal brain dysfunction. *Pediatrics,* 1972, *49,* 702-706. (c)

Conners, C. K., & Eisenberg, L. The effects of methylphenidate on symptomatology and learning in disturbed children. *American Journal of Psychiatry,* 1963, *120,* 458-469.

Conners, C. K., Eisenberg, L., & Barcai, A. Effect of dextroamphetamine on children. *Archives of General Psychiatry,* 1967, *17,* 478-485.

Conners, C. K., Eisenberg, L., & Sharpe, L. Effects of methylphenidate (Ritalin) on paired associate learning and Porteus Maze performance in emotionally disturbed children. *Journal of Consulting Psychology,* 1964, *28,* 14-22.

Conners, C. K., Goyette, C. H., Southwick, D. A., Lees, J. J., & Andrulonis, P. A. Food additives and hyperkinesis: A controlled, double-blind experiment. *Pediatrics,* 1976, *58,* 154-166.

Conners, C. K., & Rothschild, G. H. Drugs and learning in children. In J. Hellmuth (Ed.), *Learning disorders* (Vol. 3). Seattle: Special Child Publications, 1968.

Conners, C. K., Rothschild, G., Eisenberg, L., Schwartz, L. S., & Robinson, E. Dextroamphetamine sulfate in children with learning disorders. *Archives of General Psychiatry,* 1969, *21,* 182-190.

Conners, C. K., Taylor, E., Meo, G., Kurtz, M. A., & Fournier, M. Magnesium pemoline and dextroamphetamine: A controlled study in children with minimal brain dysfunction. *Psychopharmacologia,* 1972, *26,* 321-336.

Conrad, W. G., Dworkin, E. S., Shai, A., & Tobiessen, J. E. Effects of amphetamine therapy and prescriptive tutoring on the behavior and achievement of lower class hyperactive children. *Journal of Learning Disabilities,* 1971, *4,* 45-53.

Crowe, P. Aspects of body image in children with the symptoms of hyperkinesis (Doctoral dissertation, George Washington University, 1972). *Dissertation Abstracts International,* 1972, *33*(4), 1785B.

Davids, A. An objective instrument for assessing hyperkinesis in children. *Journal of Learning Disabilities,* 1971, *4,* 499-501.

Davis, K. V., Sprague, R. L., & Werry, J. S. Stereotyped behavior and activity level in severe retardates: The effect of drugs. *American Journal of Mental Deficiency,* 1969, *72,* 721-727.

Doubros, S., & Daniels, G. An experimental approach to the reduction of overactive behavior. *Behavior Research and Therapy,* 1966, *4,* 251-258.

Douglas, V. I. Stop, look and listen: The problem of sustained attention and impulse control in hyperactive and normal children. *Canadian Journal of Behavioural Science,* 1972, *4,* 259-282.

Dykman, R. A., Ackerman, P. T., Peters, J. E., & McGrew, J. Psychological tests. In C. K. Conners (Ed.), *Clinical use of stimulant drugs in children.* Amsterdam: Excerpta Medica, 1974.

Ellis, M. J., Witt, P. A., Reynolds, R., & Sprague, R. L. Methylphenidate and the activity of hyperactives in the informal setting. *Child Development,* 1974, *45,* 217-220.

Ellis, N. R., & Pryer, R. Quantification of gross bodily activity in children with severe neuropathology. *American Journal of Mental Deficiency,* 1959, *63,* 1034-1037.

Epstein, L. C., Lasagna, L., Conners, C. K., & Rodriguez, A. Correlation of dextroamphetamine excretion and drug response in hyperkinetic children. *Journal of Nervous and Mental Disease,* 1968, *145,* 136-146.

Finnerty, R. J., Soltys, J. J., & Cole, J. O. The use of D-amphetamine with hyperactive children. *Psychopharmacologia,* 1971, *21,* 302-308.

Foshee, J. Studies in activity level: I. Simple and complex task performance in defectives. *American Journal of Mental Deficiency,* 1958, *62,* 882-886.

Greenberg, L. M., Deem, M. A., & McMahon, S. Effects of dextroamphetamine, chlorpromazine and hydroxyzine on behavior and performance in hyperactive children. *American Journal of Psychiatry,* 1972, *129,* 532-539.

Grinspoon, L., & Singer, S. Amphetamines in the treatment of hyperkinetic children. *Harvard Educational Review,* 1973, *43,* 515-555.

Herron, R., & Ramsden, R. Continuous monitoring of overt human body movement by radio telementry: A brief review. *Perceptual and Motor Skills,* 1967, *24,* 1303-1308.

Horn, J. L. Review of Porteus Maze Test. In O. K. Buros (Ed.), *The seventh mental measurements yearbook.* Highland Park, N.J.: The Gryphon Press, 1972.

Jacobs, N. T. A comparison of hyperactive and normal boys in terms of reaction time, motor time, and decision making time under conditions of increasing task complexity (Doctoral dissertation, University of California, Los Angeles, 1972). *Dissertation Abstracts International,* 1972, *33*(3-A), 1045.

Johnson, C. Hyperactivity and the machine: The actometer. *Child Development*, 1971, *42*, 2105-2110.

Kagan, J., Rosman, B., Day, D., Albert, J., & Phillips, W. Information processing in the child: Significance of analytic and reflective attitudes. *Psychological Monographs*, 1964, *78* (Whole No. 578).

Keogh, B. K., Wetter, J., McGinty, A., & Donlon, G. Functional analysis of WISC performance of learning disordered, hyperactive, and mentally retarded boys. *Psychology in the Schools*, 1973, *10*, 178-180.

Knights, R. M., & Hinton, G. G. The effects of methylphenidate (Ritalin) on the motor skills and behavior of children with learning problems. *Journal of Nervous and Mental Disease*, 1969, *148*, 643-653.

Kupietz, S., Bialer, I., & Winsberg, H. G. A behavior rating scale for assessing improvement in behaviorally deviant children: A preliminary investigation. *American Journal of Psychiatry*, 1972, *128*, 1432-1436.

Lambert, N. M., Windmiller, M., Sandoval, J., & Moore, B. Hyperactive children and the efficacy of psychoactive drugs as a treatment intervention. *Journal of Orthopsychiatry*, 1976, *46*, 335-352.

Lee, D., & Hutt, C. A play-room designed for filming children: A note. *Journal of Child Psychology and Psychiatry*, 1964, *5*, 263-265.

Mann, L. Review of Marianne Frostig developmental test of visual perception. In O. K. Buros (Ed.), *The seventh mental measurements yearbook*. Highland Park, N.J.: The Gryphon Press, 1972.

McFarland, J. N., Peacock, L. J., & Watson, J. A. Mental retardation and activity level in rats and children. *American Journal of Mental Deficiency*, 1966, *71*, 381-386.

Mendelson, W., Johnson, N., & Stewart, M. Hyperactive children as teenagers: A follow-up study. *Journal of Nervous and Mental Disease*, 1971, *153*, 273-279.

Menkes, M., Rowe, J., & Menkes, J. A 25 year follow-up study on the hyperkinetic child with minimal brain dysfunction. *Pediatrics*, 1967, *39*, 393-399.

Millichap, J. G., & Boldrey, E. E. Studies in hyperkinetic behavior II. Laboratory and clinical evaluations of drug treatments. *Neurology*, 1967, *17*, 467-471.

Millichap, J. G., & Fowler, G. W. Treatment of "minimal brain dysfunction" syndrome. *Pediatric Clinics of North America*, 1967, *14*, 767-776.

Millichap, J. G., Aymat, F., Sturgis, L. H., Larsen, K. W., & Egan, R. A. Hyperkinetic behavior. *American Journal of Diseases of Children*, 1968, *116*, 235-244.

Millichap, J. G., & Johnson, F. H. Methylphenidate in hyperkinetic behavior: Relation of response to degree of activity and brain damage. In C. K. Conners (Ed.), *Clinical use of stimulant drugs in children*. Amsterdam: Excerpta Medica, 1974.

Mischel, W. *Personality and assessment*. New York: Wiley, 1968.

Page, J. G., Bernstein, J. E., Janicki, R. S., & Michelli, F. A. A multi-clinic trial of pemoline in childhood hyperkinesis. In C. K. Conners (Ed.), *Clinical use of stimulant drugs in children*. Amsterdam: Excerpta Medica, 1974.

Palkes, H., & Stewart, M. Intellectual ability and performance of hyperactive children. *American Journal of Orthopsychiatry*, 1972, *42*, 35-39.

Palkes, H., Stewart, M., & Kahana, B. Maze performance of hyperactive boys after training in self-directed verbal commands. *Child Development*, 1968, *39*, 817-826.

Patterson, G., James, R., Whittier, J., & Wright, M. A behavior modification technique for the hyperactive child. *Behaviour Research and Therapy*, 1965, *2*, 217-226.

Peterson, D. R. Behavior problems of middle childhood. *Journal of Consulting Psychology*, 1961, *25*, 205-209.

Quay, H. C., & Peterson, D. R. *Manual for the behavior problem checklist*. Champaign, Ill.: University of Illinois, Children's Research Center, 1967.

Rapoport, J., Abramson, A., Alexander, D., & Lott, I. Playroom observation of

hyperactive children on medication. *Journal of Child Psychiatry*, 1971, *10*, 524-534.

Rapoport, J., Quinn, P., Bradbard, G., Riddle, K., & Brooks, E. Imipramine and methylphenidate treatments in hyperactive boys. *Archives of General Psychiatry*, 1974, *30*, 789-793.

Routh, D. K. The clinical significance of open field activity in children. *Pediatric Psychology*, 1975, *3*, 3-8.

Routh, D. K., Schroeder, C. S., & O'Tauma, L. A. Development of activity level in children. *Developmental Psychology*, 1974, *10*, 163-168.

Sandoval, J., Lambert, N. M., & Yandell, G. W. Current medical practice with hyperactive children. *American Journal of Orthopsychiatry*, 1976, *46*, 323-334.

Schleifer, M., Weiss, G., Cohen, N., Elman, M., Cvjic, H., & Kruger, E. Hyperactivity in preschoolers and the effect of methylphenidate. *American Journal of Orthopsychiatry*, 1975, *45*, 38-50.

Schulman, J., & Reisman, J. An objective measure of hyperactivity. *American Journal of Mental Deficiency*, 1959, *64*, 455-546.

Sleator, E. K., & Sprague, R. L. Dose effects of stimulants in hyperkinetic children. *Psychopharmacology Bulletin*, 1974, *10*, 29-33.

Sleator, E., Von Neumann, A., & Sprague, R. Hyperactive children, a continuous long-term placebo controlled follow-up. *Journal of the American Medical Association*, 1974, *229*, 316-317.

Sprague, R. L., Barnes, K., & Werry, J. Methylphenidate and thioridazine: Learning reaction time, activity and classroom behavior in disturbed children. *American Journal of Orthopsychiatry*, 1970, *40*, 615-628.

Sprague, R. L., Christensen, D. E., & Werry, J. S. Experimental psychology and stimulant drugs. In C. K. Conners (Ed.), *Clinical use of stimulant drugs in children*. Amsterdam: Excerpta Medica, 1974.

Sprague, R. L., & Sleator, E. K. Effects of psychopharmacologic agents on learning disorders. *Pediatric Clinics of North America*, 1973, *20*, 719-735.

Spring, C., Blunden, D., Greenberg, L. M., & Yellin, A. M. Validity and norms of a hyperactivity rating scale. *Journal of Special Education*, in press, 1977.

Spring, C., Greenberg, L., Scott, J., & Hopwood, J. Reaction time and effect of Ritalin on children with learning problems. *Perceptual and Motor Skills*, 1973, *36*, 75-82.

Spring, C., Yellin, A. M., & Greenberg, L. Effects of imipramine and methylphenidate on perceptual-motor performance of hyperactive children. *Perceptual and Motor Skills*, 1976, *43*, 459-470.

Sroufe, L. A. Drug treatment of children with behavior problems. In F. Horowitz (Ed.), *Review of child development research* (Vol. 4). Chicago: University of Chicago Press, 1975.

Steinberg, G. C., Troshinsky, C., & Steinberg, H. R. Dextroamphetamine—responsive behavior disorder in school children. *American Journal of Psychiatry*, 1971, *128*, 174-179.

Stewart, M. A., & Haller, I. P. Letters to the editor: Hyperactive children. *Journal of the American Medical Association*, 1975, *231*, 134-135.

Sykes, D., Douglas, V., & Morgenstern, G. The effect of methylphenidate (Ritalin) on sustained attention in hyperactive children. *Psychopharmacologia* (Berlin), 1972, *25*, 262-274.

Sykes, D., Douglas, V., & Morgenstern, G. Sustained attention in hyperactive children. *Journal of Child Psychology and Psychiatry*, 1973, *14*, 213-220.

Sykes, D., Douglas, V., Weiss, G., & Minde, K. Attention in hyperactive children and the effect of methylphenidate (Ritalin). *Journal of Child Psychology and Psychiatry*, 1971, *12*, 129-139.

Victor, J. B., Halverson, C. F., Inoff, G., & Buczkowski, H. J. Objective behavior measures of first- and second-grade boys' free play and teachers' ratings on a behavior problem checklist. *Psychology in the Schools*, 1973, *10*, 439-443.

Walzer, J., Hoffman, S. P., Polizoes, P., & Engelhardt, D. M. Outpatient treatment of hyperactive school children with imipramine. *American Journal of Psychiatry*, 1974, *131*, 587-591.

Werry, J. S. Developmental hyperactivity. *Pediatric Clinics of North America*, 1968, *15*, 581-599.

Werry, J., & Quay, H. Observing the classroom behavior of elementary school children. *Exceptional Children*, 1969, *35*, 461-470.

Werry, J., Weiss, G., Douglas, V., & Martin, J. Studies on the hyperactive child III. The effect of chlorpromazine upon behaviour and learning ability. *Journal of American Academy of Child Psychiatry*, 1966, *5*, 292-312.

Whitehead, P. L., & Clark, L. D. Effect of lithium carbonate, placebo, and thioidazine on hyperactive children. *American Journal of Psychiatry*, 1970, *127*, 824-825.

Winsberg, B. G., Bialer, I., Kupietz, S., & Tobias, J. Effects of imipramine and dextroamphetamine on behavior of neuropsychiatrically impaired children. *American Journal of Psychiatry*, 1972, *128*, 1425-1431.

Witkin, H. A. The perception of the upright. *Scientific American*, 1959, *200*, 50-56.

THO

See page 265

Stop, look and listen: The problem of sustained attention and impulse control in hyperactive and normal children*,†

VIRGINIA I. DOUGLAS
McGill University and the Montreal Children's Hospital

ABSTRACT

A research programme was undertaken to investigate the specific disabilities of hyperactive children and to evaluate the effectiveness of one of the stimulant drugs on the measures found to differentiate between hyperactives and normals. It is argued that a core group of symptoms involving inability to sustain attention and to control impulsivity can account for most of the deficits found in the hyperactive group. It also appears that the stimulants exert their main effect on these deficits. Correlational and factor analytic studies suggest that the same constellation of abilities underlies the behaviour of normal children in several areas of cognitive and social functioning.

Over the past several years, together with psychiatric colleagues at the Montreal Children's Hospital,[1] my students and I have been studying children whose presenting symptoms have led to a diagnosis of "hyperactivity" or "hyperkinesis." Typically, the parents and teachers of these children complain that they are "always on the move," seem unable to concentrate, and are overly impulsive. Typically too, the hyperactivity has been present from infancy or very early childhood. I would like to review for you some of the things we have learned and some of the questions we have not yet been able to answer, because I believe that these children present a challenge not just for the clinical psychologist but for the developmental, educational, and physiological psychologist as well.

One of the first difficulties involved in trying to talk about the hyperactive syndrome is the problem of establishing a diagnosis. The literature has relied heavily on clinical descriptions and, to complicate matters further, these descriptions often overlap considerably with those given for

* Presidential Address, presented to the Canadian Psychological Association, St John's, Newfoundland, June 1971.

† This research was supported by Medical Research Council Grant No. MA-1577, Canada Council Research Grants Nos. 60-0604 and 68-1590, and a grant from the Ciba Pharmaceutical Company, Dorval, Quebec.

1 In the early years of the project we worked in close collaboration with Dr G. Weiss and Dr J. Werry of the Department of Psychiatry at the Montreal Children's Hospital. In more recent years Dr K. Minde and Dr G. Morgenstern, also from the Department of Psychiatry, joined our research team. I have found this a most rewarding venture in "interdisciplinary research" and I wish to express my sincere thanks to all of these colleagues.

children suffering from "minimal brain dysfunction" and "specific learning disabilities," two other diagnostic labels that have recently become popular.

We had hoped that it might be possible to find reasonably objective criteria for making the diagnosis, particularly since many techniques for measuring activity level are now available. However, as we have learned more about hyperactive children, we have come to believe that hyperactivity is only one of a constellation of critical symptoms. I hope to convince you that the inability of these children to sustain attention and to keep impulsive responding under control may be even more important symptoms. Other investigators (Cromwell, Baumeister, & Hawkins, 1963; Werry & Sprague, 1970) have also expressed doubt about relying too heavily on simple measures of activity level. It appears that it may be just as important to consider the quality of the hyperactive child's behaviour as its quantity. Because of his short attention span, he tends to flit from one goal to another. As a result, his behaviour is often fragmented and disorganized and these qualities may contribute to the impression of excessive activity. Thus, we decided that, at least in our early work, we would rely on the reports of parents and teachers and the judgment of experienced clinicians in choosing the subjects to include in our studies.

We have tried to keep our samples as homogenous as possible. In each case, hyperactivity had to be the major presenting complaint and the hyperactivity had to have been present from very early childhood. Both the children's parents and teachers had to report that it was excessive enough to be interfering seriously with adjustment. Children who were retarded were excluded, as were those diagnosed as neurotic or psychotic. We also ruled out children who had histories or clear symptoms of brain damage. Because brain damaged children are often hyperactive (Rosenfeld & Bradley, 1948; Burks, 1960; Levy, 1959; Strauss & Lehtinen, 1947), it has frequently been assumed that hyperactive children must have some kind of subtle brain injury. As we shall see, however, there has been considerable controversy over this assumption. Thus, we felt that it was important to check it out with a group of children who were free of the more obvious signs of brain injury.

Our first goal was to define as accurately as possible the specific disabilities that characterize the hyperactive child. Although a large number of cognitive and learning deficits have been attributed to them, including difficulties with reading, spelling, arithmetic, visual motor problems, and deficiencies in conceptual skills (Burks, 1960; Clements & Peters, 1962; Laufer, Denhoff, & Solomons, 1957; Rosenfeld & Bradley, 1948), there have been disappointingly few studies in which their performance has been compared with that of a matched group of normal controls. As a

result, a kind of clinical folklore has sprung up about the disabilities that are peculiar to hyperactive children and there has been no way to separate fact from fiction.

Another problem has arisen from the kinds of intellectual, cognitive, and personality measures that have been used. Most of the clinical instruments employed by investigators in this area tap highly complex functions which depend upon several poorly defined abilities; many of these measures also have a good deal of error variance. This is true, for example, of the subtests of the Wechsler Intelligence Scale for Children, one of the tests which has been most popular with clinicians (Cohen, 1959). Unfortunately, there is no easy solution to this problem. Our own approach has been highly eclectic. We have chosen tests commonly used by clinical psychologists and we have also borrowed extensively from the batteries that have recently been developed by educational psychologists and special educators in order to diagnose specific learning disabilities. Some of our measures are drawn from the cognitive and personality areas within developmental psychology and others are relatively well-known techniques from the experimental laboratory. We have also gone to the children's schools to find out more about the kinds of problems they are experiencing there. What I would like to do is take you along on an abbreviated version of the search that we have carried out thus far. I will begin with the more global measures and then move on to describe some of our attempts to study these children under more controlled laboratory conditions. Each study that I shall report on has involved 20–50 hyperactive children and an equal number of controls matched with the hyperactives on age, sex, overall I.Q. and socio-economic status. In each study the subjects ranged in age from 6 to 14 years and I.Q.s ranged from approximately 80 to 125 with a mean of approximately 100. It is also important to note that 80 to 90 per cent of the hyperactive children referred to us have been males.

INFORMATION OBTAINED FROM SCHOOLS

One valuable source of information on our children has been their school records and the reports of their teachers. Two honours students, Doba Lewin and Henry Laviguer, analysed data from teachers' ratings and report cards and from achievement and intelligence tests administered in the school. They found a high incidence of failed grades among the hyperactives. By the time they were 12 years of age, 70 per cent of them had failed one year, and 20 per cent had failed two years. The hyperactive children also had significantly lower grades than the controls on almost all academic subjects. The teachers' ratings which showed the highest differences were concerned mainly with frustration tolerance, concentration, and ability to organize one's own activities. On group administered achievement tests, the hyperactives did significantly more poorly than the normal controls on the Stanford Arithmetic Test and the reading speed score of the Gates Reading Test. They did not do more poorly than the controls, however, on either the comprehension or vocabulary scores of the Reading Test. They also received lower I.Q. scores than the controls on group administered intelligence tests; their scores on these group tests were significantly lower than those obtained on individual intelligence tests administered at the hospital.

Another group of students has taken on the task of collecting observational data on the children's behaviour in the classroom. Sandra Witelson did most of the original work involved in developing the observation schedule and the study of hyperactive-normal comparisons was done by Christine Bradley. She studied two groups of hyperactive children – the mean age of the younger group was 7 years, 9 months and that of the older group was 12 years, 8 months. She obtained the most consistent differences between hyperactives and normals on a coding category that included "purposive (or goal-directed) behaviour not related to the classroom activity." It is important to note that we are not talking here about random or aimless behaviour. The problem seems to be, rather, that the child's goals and the teacher's goals differ. We also observed various kinds of "fidgeting" behaviour but our findings on this category have not differentiated reliably between hyperactives and controls. The differences between the younger and the older hyperactives were also interesting. The younger ones moved around the classroom and vocalized more than their controls. They also showed more disorderly behaviour toward the teacher and attracted more attention from her. The behaviour of the older group on the other hand, was less disrupting: although they also engaged in more purposeful behaviour that was not the classroom activity, in their case this behaviour took place while they were sitting in their seats. They were frequently observed, for example, to be working on the wrong assignment or playing with a toy.

CLINICAL AND LEARNING DISABILITY TESTS

I would like to review next the results we obtained from the clinical and learning disability battery. From time to time the clinical literature has reported that hyperactive children have difficulty on particular subtests of the Wechsler Intelligence Scale for Children (Burks, 1960; Clements & Peters, 1962). However, we found no consistent subtest pattern nor did we find the differences between the verbal and performance sections of the

test which have sometimes been described. We have repeatedly found, however, that the hyperactive children show more variability from subtest to subtest. As we shall see later, this kind of variability also typifies their performance on other tasks. The hyperactives have also consistently scored lower than the normals on the Goodenough Harris Draw-a-Person Test, the Bender Visual-Motor Gestalt Test, and the Lincoln-Oseretsky Schedule of Motor Development. They had some difficulty, too, with the Frostig Developmental Test of Visual Perception; this showed up in a significantly lower overall score and a poorer score on the eye-motor co-ordination subtest. This last group of findings seems to suggest that the hyperactive children have unusual difficulties on visual motor tasks and tasks requiring fine and gross motor co-ordination. Similar reports appear in the literature (Clements & Peters, 1962; Laufer, et al., 1957). I would like to point out, however, that besides testing visual-motor ability, all of these tests require the child to perform a task with care and concentration; the Bender is also thought to require an analytic approach. I will return to this point later.

It is important to note that on the remainder of the rather extensive battery used, there were almost no differences between the hyperactive and control subjects. This included an individually administered reading test, a test tapping several different abilities in the language area, and tests of such abilities as auditory discrimination, right-left discrimination and short term memory. In all, there were 41 measures on which we found no differences between the hyperactive and normal children.

AUTOMATED CONCEPT LEARNING TASK

In her doctoral dissertation, Vaira Freibergs (Freibergs, 1965; Freibergs & Douglas, 1969) focused on concept learning in these children. The clinical literature had suggested that they would have difficulty functioning at a conceptual level (Burks, 1960; Rosenfeld & Bradley, 1948). Freibergs used a concept learning apparatus and several concept problems developed by Sonia Osler at Johns Hopkins University (Osler & Fivel, 1961). The machine delivers two pictures at a time. One is an exemplar of a concept such as "flower" or "bird" or a number concept such as "two"; the other is a non-exemplar of the concept. In the case of the bird or flower concept, for example, the non-exemplar may be any one of a variety of objects such as a house, an animal or an aeroplane; in the case of the "two" concept it is another number. The child is told that if he looks at the pictures very carefully he will find that there is something in them, like an idea, which will tell him which one to choose to get a marble as often as possible. He makes his choice by pressing a lever. It is thus possible to observe, step by step, the child's attempts to solve a new problem. A record is kept of the number of errors he makes and the number of trials to criterion.

Performance on these tasks was studied under three conditions: a continuous reinforcement condition in which every correct response was rewarded, a 50 per cent partial reinforcement condition in which every other correct response was rewarded and a delay condition in which the interval between presentation of stimuli was increased from 4 to 8 seconds. Under continuous reinforcement, the hyperactive and normal children reached criterion equally quickly, although on the first problem there were differences in the shapes of their learning curves which suggest possible differences between the groups during the presolution phase. The performance of the hyperactives was highly erratic, showing wide fluctuations from their own mean performance. By the second problem, however, this kind of fluctuation seemed to disappear. In both Freiberg's study and a later study by Sheila Macklin, they also showed excellent transfer from problem to problem. We were struck by the way that learning in the hyperactive children seemed to thrive under the circumstances provided by this teaching device in the continuous reinforcement condition. The stimuli were colourful and interesting and they remained on the screen until the subject made his choice. Perhaps even more important, the machine provided immediate feedback and reinforcement for correct responses.

Freibergs also investigated the children's ability to reverse concepts. After the subjects reached criterion the machine was switched so as to deliver a marble when the *non*-exemplar of the concept appeared. There were no differences between the hyperactive and control children in their ability to reverse. Thus, on this task, we find no evidence to support reports in the clinical literature that hyperactive children tend to perseverate. It should be noted, however, that on this particular task the non-exemplars of the concept provide more variety and novelty than the exemplars; we have other data which suggest that the hyperactives may have a greater pull towards novelty. Doubling the intertrial interval from 4 to 8 seconds did not produce significant differences between the hyperactive and normal subjects although there was a tendency for the hyperactives to do less well.

The most striking difference between the two groups appeared under the 50 per cent partial reinforcement schedule. Although both hyperactives and normals had more difficulty under this condition, the performance of the hyperactive children was much more severely impaired. Sixty-five per cent of them failed to reach a solution within 300 learning trials. There are several possibilities which might explain why the hyperactives did so badly when the ratio of reinforcement was reduced. Unfortunately, it is

difficult to sort out the effects of motivational, attentional, and information feedback variables. However, Penny Parry is currently making an attempt at accomplishing this as part of her doctoral thesis. We are hoping that what she learns will provide some valuable insights that may help in training these children.

STUDIES OF ATTENTION

Donald Sykes's doctoral dissertation (Sykes, 1969; Sykes, Douglas, Weiss, and Minde, 1971; Sykes, Douglas, and Morgenstern, 1972) was designed to focus on some critical aspects of attention and followed up earlier work done by Edgar Zurif for his master's degree. There are numerous reports in the literature which suggest that attention is impaired in hyperactive children. Little has been done, however, to define empirically the kinds of attentional problems they demonstrate. Sykes's studies were designed to point up some of the factors that might affect their performance in learning situations.

In one task of choice-reaction-time the child was required to respond to specific stimuli but was warned each time before the stimuli appeared: on this task the hyperactive children had no difficulty. A second task, a serial reaction task, also required responses to particular stimuli but in this case the child had to continue responding over a prolonged period of time (15 minutes). The task was self-paced, that is, each stimulus appeared only after the child had made his response to the previous stimulus. Here the performance of the hyperactive children was somewhat impaired: they made more incorrect responses (responses to wrong stimuli) than the controls but they apparently worked quickly enough to make a similar number of correct responses.

The attention task on which the hyperactive children had most difficulty was a continuous performance task. This was a vigilance task developed by Mirsky and Rosvold (1963). It was experimenter-paced and the subject was required to perform over a 15-minute interval. He had to respond to particular stimuli, for example, the letter X when it was preceded by the letter A. The stimuli arrived automatically on a screen or through earphones. On both the visual and auditory forms of this task the hyperactive children identified fewer of the correct stimuli and also responded more frequently to incorrect ones; several of these latter errors appeared to be of an impulsive nature. The performance of the hyperactives also deteriorated more seriously over time than that of the normals. This deterioration seemed to be accompanied by increased motor restlessness which was measured by a stabilimetric cushion (Sprague and Toppe, 1966) attached to the child's chair.

Sykes also looked at the effect of piping intermittent white noise (80 decibels) into the experimental room at frequent random intervals while the subjects were performing on the continuous performance test. This attempt to distract the children did not affect the hyperactive subjects more than the normals. An attempt was also made to introduce conflicting cues on the choice reaction time test; in one series of trials the coloured background of the stimuli appearing on the screen was discrepant with the colour of the background of the target stimuli. This procedure also failed to affect the hyperactives differentially.

Nancy Cohen was also interested in attentional processes in children and her dissertation involved the monitoring of autonomic components of the orienting response and its habituation while the children were responding to different task demands (Cohen, 1970; Cohen and Douglas, 1972). She took skin response and heart rate measures under a variety of conditions: first during a rest period, then while the child was listening to stimuli to which he was not required to respond and, finally, while he was reacting to warning and reaction signals while performing on a delayed reaction time task.

Let me give you a brief summary of her results. First, there were no differences in the tonic levels of skin conductance or of heart rate during relaxation periods. Secondly, when the Ss were simply required to sit and listen to a series of tones which were presented through earphones there were no significant differences between hyperactive and normal subjects either on tonic levels of skin conductance or on the orienting response to the first or later tones. I would like to mention, however, that Cohen has some preliminary data which suggest that there may be differences when the tones are piped into the room through a loud speaker rather than through earphones. We plan to look more carefully into these effects.

When we look at the orienting response (OR) measures during the Delayed Reaction Time Task, where the Ss were required to make active responses to discrete stimuli, significant differences between the hyperactive and normal subjects emerge. The controls exhibited a significant increase in both tonic and phasic OR measures while the hyperactives remained relatively unresponsive. Performance on the reaction time task itself was also clearly deficient in the hyperactive children; compared with controls they exhibited slower reaction times and a greater amount of variability in performance. It would appear as if, for the hyperactive children, the warning signal given at the onset of each reaction time trial did not have the intended effect of alerting the child and preparing him to respond to the reaction signal.

It is important, however, to emphasize the extreme variability in the performance of the hyperactive children on this task. If we look at their

best trials, we find that they are capable of reacting as quickly as the normals; it is the erratic nature of their performance that reduces their scores. It is probably this kind of behaviour which inspires the lament so frequently heard from teachers: "He can do it if he wants to."

Cohen also attempted to improve the hyperactives' performance by reinforcing them for performing well. She did this simply by saying "good" on every trial on which the child surpassed his own basal reaction time score. After 15 trials with reward, the child was told that the examiner would no longer say anything but that he should still try to respond as quickly as possible. The reinforcement produced faster reaction times in both hyperactives and controls. There was also a decrease in variability in performance in both groups. However, when reward was withdrawn, the reaction times of the hyperactives tended to slow more quickly than those of the controls. The autonomic records during the reward condition also show differences between the hyperactives and controls. Although tonic activity increased for both groups when reward was introduced and stayed at the increased level when it was withdrawn, the pattern for the OR to the warning signal was quite different in the two groups. For controls, OR frequency increased significantly when reward was given and when it was withdrawn the frequency was still significantly higher than in the initial non-reward period. For hyperactives, on the other hand, OR frequency remained relatively low and did not change with the altered reinforcement conditions. While the tonic response is thought to measure a general increase in alertness or non-specific activation, the phasic response is thought to render the subject more sensitive to specific incoming stimuli (Lynn, 1966). It is this more specific response that seems to have been lacking in the hyperactive group. We feel that these findings underline the importance of training hyperactive children to concentrate on the critical aspects of a learning situation: it is apparently not sufficient just to increase general motivation. Recent findings by Penny Parry have also emphasized this point. She has discovered that a high level of general, non-contingent praise can even lead to a deterioration in performance.

Cohen also recorded other irrelevant movements while the subjects were performing on her task. The hyperactive children showed a strong tendency to press and release the response button after the appropriate response had been made: they also exhibited more intense movements in the left hand simultaneous with the appropriate response to the reaction signal with the right hand. One might think that such movements would interfere with the performance of these children. However, during reward, the frequency of these responses increased along with the improved performance. Thus, there does not seem to be a simple negative relationship between efficiency in attending and the amount the child moves. This and other observations we have made on these children have led us to question the importance that some educators and investigators have attached to making them sit still in order to get them to attend.

Susan Campbell's dissertation (Campbell, 1969; Campbell, Douglas, & Morgenstern, 1971) concentrated on variables quite different from the ones already discussed. Here, interest was in the style or approach that hyperactive children typically employ in problem solving. There has been a growing body of literature in the field of developmental psychology that has revealed the pervasive influence of several different cognitive styles on the approach that individuals take to a variety of problem solving situations. Several of the styles are relatively unrelated to intellectual ability. We felt that an understanding of the cognitive styles of hyperactive children might shed some light on their learning and behavioural problems and might also suggest directions for training and educational planning. Campbell studied four of the cognitive styles but I shall concentrate on two: field dependence-independence and reflection-impulsivity.

Kagan and his associates (Kagan, Rosman, Day, Albert, & Phillips, 1964) have demonstrated that one factor contributing to differences in cognitive functioning is the child's cognitive tempo, that is, his habitual speed of decision-making in situations with high response uncertainty. The impulsive child makes his decisions too quickly and, as a result, is likely to make errors, while the performance of the reflective child is characterized by long latencies and few errors. Normal children scoring high on impulsivity measures have been found to make more errors of commission on a serial learning task and to have higher error scores on tests of inductive reasoning. They have also been observed to be more distractable, less attentive, and more physically active than their more reflective peers (Kagan, 1965; 1966; Kagan, Pearson, & Welch, 1966). To measure reflection-impulsivity, Campbell used Kagan's Matching Familiar Figures Test which consists of sets of pictures of common objects and animals. The child is shown a standard stimulus and six similar ones and is required to choose the one picture from among the six alternatives which is identical with the standard. This is not particularly easy to do – many of the possible choices are so close to being correct that the child must hold back his tendency to respond long enough to look over the pictures very carefully. The test is scored for latency to first response and number of errors.

Witkins' dimension of field dependence-independence (Witkin, Dyk, Faterson, Goodenough, & Karp, 1962) reflects individual differences in the ability to separate an item from the field in which it is embedded. This dimension is also conceptualized in terms of differences in the degree to which perception is global and diffuse or structured and analytic. The field-independent individual is better able to overcome a confusing,

embedding context when isolating figure from ground. Witkin and his associates have demonstrated that field independent boys are more emotionally independent and display better control over impulses than field dependent boys (Witkin et al., 1962). Field-independent boys are also more concerned with intellectual tasks. Campbell used the Children's Embedded Figures Test (Karp & Konstadt, 1963) as her measure of field-independence. It consists of two series of simple figures embedded in more complex designs; the score is based on the number of figures correctly located.

Campbell found that both of the cognitive style measures differentiated significantly between hyperactives and controls. On the Matching Familiar Figures Test the hyperactives had significantly shorter latencies and made more errors. On the Embedded Figures Test they isolated fewer embedded figures. Thus, they would be classified as being more impulsive and more field-dependent than the normal controls. In her discussion of these results, Campbell suggests that it may be possible to modify the cognitive styles of hyperactive children. Training could be centred on teaching the child to delay responding until he has examined the various response alternatives in a situation. One could work up gradually from a few simple alternatives to several more complex ones. It might also be possible to use materials similar to the embedded figures so as to help him learn to focus on the essential features of a situation. He could be taught organized strategies of search for discovering the figure, first with simple examples, and later with more complex ones. Successful approaches for training children with attentional and impulsivity problems have been developed by Palkes, Stewart, and Kahana (1968), Santostefano and Stayton (1967), and, more recently, by Meichenbaum and Goodman (1971). We have also begun some preliminary work with a group of highly impulsive disadvantaged children. We have used games such as "May I," and a variety of card games and interesting projects in which the child can only succeed if he keeps his impulses under control. The teacher tries to teach and demonstrate "reflective" strategies and encourages the child to verbalize them and carry them out.

Before I leave Campbell's study, I want to mention just briefly that she also looked at a cognitive style called constricted-flexible control (Klein, 1954). She used the Colour Distraction Test developed by Santostefano and Paley (1964) which requires the child quickly to name the colour of objects arranged on a card, while ignoring distracting and contradictory stimuli. There was no evidence that the performance of the hyperactive children was impaired under either the distracting or interference conditions. Thus, this is one of several times we have failed to show any significant effect of distracting stimuli on our hyperactive children. We obviously need to learn a good deal more about attentional and distractibility problems in these children. I do want to emphasize, however, that their attentional problems appear even when they are working alone in a relatively empty, sound-proof room.

EFFECTS OF THE STIMULANT DRUGS

Now that we have established a pretty clear picture of some of the disabilities of hyperactive children, I would like to turn to a consideration of how drugs effect their performance on these same tasks. I am going to concentrate here on our studies using methylphenidate (Ritalin) which is one of the stimulants. We did do some earlier work with chlorpromazine (Freibergs, Douglas, & Weiss, 1968; Werry, Weiss, Douglas, & Martin, 1966), but found no evidence that it improved cognitive functioning. In the studies that I shall report the sample size varied from 20 to 50 subjects. Most of the studies used a double blind, cross-over design. The psychiatrist was allowed to control dosage to achieve a maximum effect. The dosage of methylphenidate was gradually increased from 5 mg/day to a maximum of 100 mg/day. Mean dosage was approximately 60 mg/day. Test-retest interval was two weeks.

The drug produced some positive effects on the tests in the clinical and learning disability battery: there was a significant increase in overall I.Q. and also on a few of the other tests in the battery; however, we could see no particular pattern in the kinds of tests affected. There was also significant improvement on several of the attention measures used by Sykes (Sykes, 1969; Sykes et al., 1971; Sykes et al., 1972). On the Continuous Performance Test on which the hyperactives had shown their most serious impairment, the number of correct responses increased, the number of incorrect responses decreased, and there was considerably less deterioration in performance over time when the children were receiving the drug. On the Serial Reaction Time Task, there was improvement on both correct and incorrect response measures; on the Choice Reaction Time Task, reaction time decreased. Thus, the drug produced positive changes on the measures on which the hyperactives had been most deficient but it also produced significant improvement on some measures on which there had been no significant differences between hyperactives and controls. On Cohen's delayed reaction time task (Cohen, 1970; Cohen, Douglas, & Morgenstern, 1971) the mean reaction times were faster and less variable and task-irrelevant motor responses were less frequent when the hyperactives were receiving the drug. On Campbell's cognitive style measures (Campbell, 1969; Campbell et al., 1971) the hyperactives became significantly less impulsive while on methylphenidate. This was reflected in

longer reaction times and fewer errors on the Matching Familiar Figures Test. The differences on the Embedded Figures Test of Field Dependence-Independence did not reach significance. Taken together, these results suggest to us that methylphenidate exerts its main effect by helping the hyperactive child sustain attention and control impulsivity. Similar results are being reported from several other laboratories (e.g., Conners & Rothschild, 1968; Knights & Hinton, 1969; Sprague, Barnes, & Werry, 1970).

SPECULATIONS ABOUT AETIOLOGY

There has been a good deal of physiological speculation about how the stimulants produce their apparently "paradoxical" effect and these theories, in turn, are tied to further speculations about the aetiology of the hyperactive disorder. Many clinicians and investigators believe that the hyperactive child has undergone some kind of brain damage; birth injuries appear to be the most frequently cited explanation (Laufer & Denhoff, 1957; Martin, 1967; Gross & Wilson, 1964; Rosenfeld & Bradley, 1948; Levy, 1959). It is, however, commonly recognized that no evidence of brain damage can be found in many of these children and so the more ambiguous notion of "brain dysfunction" has become popular (Clements & Peters, 1962; Stevens, Boydstun, Dykman, Peters, & Sinton, 1967). Recently, several writers have suggested the possibility of a biochemical defect (Shetty, 1971; Stewart, 1970; Silver, 1971); norepinephrine is the chemical most mentioned, partly because its release is thought to be stimulated by the amphetamines (Carr & Moore, 1969; Stein & Wise, 1969). Several of the investigators who have adopted a biochemical explanation believe that hyperactivity is a hereditary trait (Stewart, 1970; Silver, 1971). The fact that it is so much more common in males than in females suggests the possibility that it is a sex-linked or sex-influenced character; males, however, are also known to be more vulnerable to brain injuries.

There has been a good deal of theorizing about the nature of the brain mechanisms underlying the disorder. Among the possibilities that have been mentioned are the failure of some essential inhibitory control or filtering mechanism (Laufer, et al., 1957) and an imbalance between cortical and subcortical structures which results in the cortex having insufficient control over subcortical centres (Knobel, Wolman, & Mason, 1959). The reticular activating system plays an important role in several of these theories. Many of the theorists also think of the hyperactive child as suffering from either an abnormally high or abnormally low level of physiological arousal (Laufer, et al., 1957; Werry, Sprague, Weiss, & Minde, 1969; Stewart, 1970; Satterfield & Dawson, 1971). These beliefs are based on the children's high level of behavioural activity, the fact that they appear to be searching for sensory and kinesthetic input and the effectiveness of the stimulant drugs.

Our own research has added little that is definitive to these speculations. Members of our team have failed to find substantial evidence of brain damage in our children's birth histories, electroencephalograms or neurological examinations (Werry, Weiss, & Douglas, 1964; Werry, Weiss, Dogan, Minde, & Douglas, 1969; Minde, Webb, & Sykes, 1968), although, like the psychological tests, the neurological examinations have shown them to be poorly co-ordinated. It is important to remember, as I mentioned earlier, that we excluded from our samples children with clear evidence of brain damage: however no initial screening was done for more subtle signs. These studies have clearly emphasized to us the importance of proper control groups and follow-up studies in work of this kind; a surprisingly large number of our controls for example, had "abnormal" EEGS. We were also struck by the poor reliability of the electroencephalographic data over time. It should be noted, however, that we have not worked with the more sophisticated EEG techniques such as evoked potentials and contingent negative variation.

Although I am a little reluctant to add further to the rather loose physiologizing that has been going on, I would like to mention that Penny Parry has been experimenting with two tests which have been found to differentiate patients who have undergone frontal lobe surgery from patients who have lesions in other areas (Milner, 1963; Milner, 1964; Porteus, 1959). These are the Wisconsin Card Sorting Test which requires the ability to switch set and the Porteus Mazes which measure the ability to plan and follow rules. The hyperactives are doing very poorly on both of these tests, though at this point we are not sure that the reasons are always the same as those put forward by Milner (1964) to explain the poor performance of her frontal-lobe patients. I can only say at this time it is our impression that attentional and impulsivity problems contribute greatly to their performance. So far as "physiological arousal" is concerned, I am well aware of the complexities involved in using this term. However, for what it is worth, let me remind you that Cohen found no differences between resting levels of hyperactives and normals on either skin conductance or heart rate measures. Differences between the two groups showed up only on orienting response measures when the children were required to attend and respond to stimuli. However, this matter is far from settled. Although the findings of some other investigators (Stevens, Boydstun, Ackerman, & Dykman, 1968; Boydstun, Ackerman, Stevens, Clements, Peters, & Dykman, 1968) have tended to agree with our results, Satterfield & Dawson (1971) have reported first abnormally low and more recently (personal communication) abnormally high arousal levels in their hyperactive

groups. They have suggested that their conflicting findings may be explained by differences in the laboratory conditions in their two studies. As I mentioned earlier, physiological recordings in these children seem to be unusually sensitive to relatively minor changes in methodology. I suspect that this is due to the attentional problems I have described and also, perhaps, to a high degree of lability in both their behavioural and physiological responsivity. Their overreaction to praise which I mentioned earlier may be one example of this.

WHEN TO USE THE STIMULANTS?

It must be clear by now that although the physiological speculations about the hyperactive child raise many intriguing questions, they cannot offer any definitive answers. So far as the stimulant drugs are concerned, they seem to "work" with at least some of these children but we do not know why. This leaves the clinician open to severe ethical conflict about when the drugs should be used. Unfortunately, the question is in danger of becoming a political issue. I recently found myself embroiled in a debate with John Holt on this topic in the *New York Review of Books* (13 August, 22 October, and 3 December 1970). Not surprisingly, Holt places a good deal of blame for the problems of hyperactive children on the shoulders of unimaginative middle-class teachers; he also accuses the physicians who prescribe the drugs of "fashionable quackery." As one reads what he has to say, it becomes clear that no dull, middle-class professor ever taught him to do a literature review before allowing him to pontificate on a subject. Nevertheless, Holt and others who have voiced concern do raise a valid point. Although many of our children have expressed relief at finally being able to control themselves, I and most of my colleagues have come to believe that these drugs should be used only if the child's symptoms are extremely debilitating. Some of the stimulants are known to be addictive in adults and, though no cases of addiction have been reported by clinicians working with hyperactive children or adolescents, the possibility cannot yet be dismissed. We have also found greatly increased heart rates in some of the children in our short-term studies (Cohen & Douglas, 1972); we have no data as yet on long term effects. We find, too, that a few of our children become extremely depressed and show strangely flattened affect while they are on the drugs. Thus, we believe that the drugs should be used only after a very careful evaluation of the child's problems and if they are used, the physician should stay in close contact in order to titrate dosage and to monitor the youngster's response. It is to be hoped that, as our diagnostic techniques are sharpened, psychologists will be able to provide more help both in identifying children who may benefit from the stimulants and in evaluating their effects. As you might guess, I place most faith in methods for diagnosing attentional and impulsivity problems. However, though I believe this contribution would be a valuable one, my hope is that we will also be able to contribute to the development of training techniques which may at least diminish the need for drugs.

DOES THE HYPERACTIVE CHILD OUTGROW HIS SYMPTOMS?

I would like to look now at what we have been able to learn about the developmental history of the hyperactive disorder. As I have said, in the children we have studied, the symptoms have been present since infancy or early childhood. But what about the progress of the disorder as the child matures? There have been many clinical reports suggesting that these children "outgrow" their symptoms. If this were so, it might provide some evidence for the theories which suggest that the symptoms are a result of a "maturational lag" in the development of some critical part of the central nervous system (Lytton & Knobel, 1958; Solomons, 1965). Gabrielle Weiss has tried to stay in touch with our hyperactive children as they grow towards adolescence and we now have some data on a group of 13 to 16 year olds. There is some evidence, as there was in the school observations, that the hyperactivity has decreased; however, it is clear that impulsivity and inability to attend remain a problem. Cohen stayed on with the project after completing her degree and has run several of our tests in the longitudinal study (Cohen, Weiss, & Minde, in press). These young adolescents were still making more errors than their controls on the Continuous Performance Test, both on the visual and auditory form. On the Matching Familiar Figures Test of Reflection-Impulsivity they still reacted more quickly and they also made more errors than the controls on the Embedded Figures Test. Thus, they remained more impulsive and more field dependent. Cohen found no differences between normals and controls on the Stroop Colour-Word Interference Test. This, like Santostefano's Colour Distraction Test, which I mentioned earlier, is a measure of constricted-flexible control and is thought to measure the extent to which individuals are susceptible to cognitive interference from conflicting cues. This, then, represents another failure to prove that the hyperactives are unusually distractible. Psychiatric interviews revealed that 30 per cent of the children were described by their mothers as having *no* steady friends and 25 per cent had a history of fairly serious acting-out and anti-social behaviour (Weiss, Minde, Werry, Douglas, & Nemeth, 1971). Teachers' ratings and school reports analysed by Doba Lewin and Henry Lavigueur revealed that the children were still doing significantly worse than their controls at school, both academically and socially. Thus, these first follow-up investigations provide little hope that maturation is restoring these children to normality.

STOP, LOOK AND LISTEN!: AN UNDERLYING DIMENSION?

Let me try to summarize what we think we have learned thus far about the pattern of our hyperactive children's deficiencies. First, we should not forget that there are several areas of functioning that are relatively unimpaired in these children. Many of our subjects obtain average or even well above-average I.Q.s on standard intelligence tests, particularly if they are individually administered. As a group, they show no significant differences from normals in terms of language abilities, comprehension, or conceptual thinking. Neither is there any evidence of difficulty with short-term memory, even though we tested for this with several different kinds of materials. They appear, too, to be less disrupted by outside distractions than many of the reports in the literature would suggest; nevertheless they can be led astray by stimuli that are highly attractive to them. Although they do move more than other children, most of their behaviour seems to be directed to obvious goals, albeit their own. There is some suggestion, too, that activity level may not be the most critical aspect of the symptom picture of these children. For example, when they grow older they become less active but attentional and impulsivity problems remain.

As I looked back over our various studies, it struck me that one closely related group of characteristics can pretty well account for all of the deficiencies we have found. These youngsters are apparently unable to keep their own impulses under control in order to cope with situations in which care, concentrated attention, or organized planning are required. They tend to react with the first idea that occurs to them or to those aspects of a situation which are the most obvious or compelling. This appears to be the case whether the task requires that they work with visual or auditory stimuli and it also seems to be true in the visual-motor and kinaesthetic spheres. These same deficiencies – deficiencies which I have come to think of as the inability to "stop, look and listen" – seem also to influence the children's social behaviour. Penny Parry has shown that on a story completion task, they are unable to react realistically to a potentially frustrating situation, and, in real life, several of our older hyperactives are beginning to get into trouble with the law because of their inability to control their impulsive tendencies.

We have been struck by the degree to which our measures that tap attention, impulse control, and the ability to take an analytic approach to problems seem to go together in these children. These congruences were found in our early testing and the same *constellation of deficiencies is pres-*

ent in our follow-up study. Furthermore, methylphenidate seems to exert its effects on these same measures. Since Sykes, Cohen, and Campbell all used the same subjects in their studies we were able to run correlations among their various measures; as expected, we found rather high intercorrelations. For example, correct responses on the continuous performance test correlated −0.66 with mean delayed reaction time and also −0.66 with errors on the Matching Familiar Figures Test.

RELEVANCE OF FINDINGS TO NORMAL POPULATION

All of this made us wonder whether scores from these same tests would be intercorrelated in a normal population. We were encouraged to think that this might be so by some findings which Campbell had obtained from a group of normal children (Campbell & Douglas, 1972). Her data had revealed significant correlations between her subjects' field dependence scores and their scores on the reflection-impulsivity measures. Furthermore, both measures were significantly correlated with scores on a story completion test that gauged the children's optimism about their ability to cope with a difficult, potentially frustrating situation. These correlations were independent of the effects of both intelligence and age. We therefore decided to investigate the possibility that the attention-impulse control constellation which seemed so important in the functioning of hyperactive children might also be an important factor underlying the cognitive and social functioning of normal children. We tested this out by administering to a group of 41 eight- and nine-year-old normal boys the same tests that we had found to differentiate between hyperactive and normal children. We also included in the test battery four I.Q. measures and predicted that these would be relatively unrelated to our hypothesized dimension. Finally, we added two anxiety measures which we also expected to be independent of our attention impulse control factor.

The results were extremely interesting. The correlational matrix revealed a very clear pattern of significant intercorrelations among most of the tests which we thought would tap our underlying "stop, look and listen" dimension. Very few of the measures correlated significantly with the I.Q. measures: there were also few significant correlations with the anxiety measures. A factor analysis was then performed; a varimax rotation produced four factors. The first of these looks very much like our hypothesized dimension. The tests loading significantly on it (in order of their loadings) included: the Porteus Mazes, the Children's Embedded Figures Test of Field Dependence-Independence, teachers' ratings on a hyperactivity scale, the eye-motor co-ordination subtest of the Frostig Motor Development Schedule, aggressive responses on a story completion test, the Bender

Visual-Motor Gestalt Test, a listening task involving loudness discrimination, the Matching Familiar Figures Test of Reflection-Impulsivity, the Continuous Performance Test, and responses to a story completion test which demonstrated the child's ability to cope realistically with a frustrating event. The second factor seems clearly to be an intelligence factor. Three of the intelligence tests loaded on it as well as two of the story completion test scores. The third is probably an anxiety factor. It loads on our two anxiety measures but also on the Lincoln-Oseretsky Schedule of Motor Development and the Bender Visual-Motor Gestalt Test.

These results have been obtained only recently and we have not had much chance to follow up their implications. At the moment we are inclined to think that our "Stop, Look and Listen" dimension is not unique to children diagnosed as hyperactive but is, rather, an important factor influencing the capacity of all children to cope effectively with a wide range of situations. We are interested in learning more about what the distributions of normal and hyperactive children look like on the various tests that define this factor. At the moment, I can only say that the distributions of scores of the "normal" sample on several of the tests seem to reveal considerable skewness on the "bad" end. It is as if something is responsible for producing more children who do badly on these tests than one would expect by chance.

Perhaps it may seem that I am pushing my "Stop, Look and Listen" dimension too far. However, before I conclude, I would like to try to stretch its implications to one more behavioural area – the area of moral behaviour. Michael Schleifer's doctoral dissertation (Schleifer, 1971; Schleifer & Douglas, in press; Schleifer & Douglas, unpublished manuscript) dealt with moral judgments in young children. His measures are based on Piaget's approach to moral development which parallels closely his theorizing about cognitive development. Piaget believes that the child's early moral judgments are limited by his cognitive capacities. For example, he tends to be very much influenced by the behaviour of authority figures because he sees moral rules as fixed laws. It is only later that he realizes that rules are made by parents who are not necessarily infallible. The young child is also limited by his "egocentrism." He is unable to take the viewpoint of another person and thus to take that person's intentions into account when he is making judgments about the goodness or badness of the person's acts. Piaget speaks also of "syncretism" a term which refers to the child's reacting globally to a situation rather than analysing its elements; the related concept of "centration" refers to the tendency to focus on some striking but superficial aspect of a phenomenon. By now it can be seen that we are approaching morality from the point of view of the ego rather than the superego. I hope that there can be detected in the above descriptions certain similarities with the kinds of abilities I have already discussed.

Schleifer took on the job of seeing whether there was, in fact, a relationship between our measures and the stages of moral development posited by Piaget. His test instruments were stories very similar to those used by Piaget and also, for his younger subjects, films which depicted situations similar to the ones in the stories. The stories and films portray situations in which a child commits an act such as breaking something. The subject is questioned about the goodness or badness of the child in the story and is asked to give the reasons behind his judgment. For example, there is one film in which a boy breaks several glasses while trying to help his mother with the dishes. In another film, a boy breaks a single glass while trying to steal cookies. Young children are likely to consider the first boy to be the really bad one because they fail to take intentions into account and also because they tend to be overly influenced by the amount of damage done. The children's responses to a series of stories (or films) were used to yield a "moral maturity score." The other measures included tests of reflection-impulsivity and field dependence-independence, as well as teachers' ratings on several behavioural scales.

Now for a brief summary of Schleifer's findings. In three different samples, significant correlations were obtained between moral judgment, field-dependence-independence and reflection-impulsivity scores. Most of the correlations ranged between 0.4 and 0.6. It is also important to note that none of the three measures correlated significantly with intelligence as measured by the Peabody Test of Verbal Intelligence. The data from teachers' ratings also lend support to our previous findings. In one sample, correlations between teachers' ratings of the child's propensity towards aggressive behaviour and the measures of moral judgment, field independence and reflectivity were negative and significant. In another sample, children high on the morality measure were also rated by their teachers as being more attentive and less impulsive.

Lest it is beginning to appear that I have got myself out on a limb, let me just say that I have some company. We recently discovered an article on "Conscience and Attentional Processes" by Grim, Kohlberg, and White (1968) in which the authors indulge in reasoning very similar to ours. They report on the chance meeting of two studies which happened to be going on in the same school. One group of investigators was studying attentional processes in children, using a reaction time task and associated galvanic skin response measures; their methods and measures were very similar to the ones used by Cohen in her dissertation. The other group of investigators was studying cheating behaviour, using situational tests of cheating and teachers' ratings of the children's ability to resist temptation.

As we would have predicted, there were several significant correlations between the measures obtained from the two studies. In discussing their results, the authors suggest that children who can maintain stable attention are also able to resist the quick, effortless solution obtained by cheating. The authors themselves are sufficiently honest and morally mature, however, to admit that they are not the first psychologists to see the theoretical connection between attentional processes and moral behaviour. Perhaps you will recognize the author of the following quotation: "If a brief definition of ideal or moral action were required, none better would fit the appearance than this: it is action in the line of greatest resistance ... We reach the heart of volition when we ask by what process it is that the thought of any given object comes to prevail stable in the mind. Attention with effort is all that any case of volition implies. The essential achievement of will is to *attend* to a difficult object and hold it fast before the mind."

I am sure you will have guessed that the author is William James (1890). So, like many psychologists pushing their theories, I have discovered that James got there before me!

RÉSUMÉ

Exposé d'un programme de recherche pour étudier les déficits spécifiques des enfants hyperactifs et l'efficacité des agents psychoanaleptiques sur les mesures prises en vue de distinguer normaux et hyperactifs. Il est allégué qu'un noyau de symptômes centrés sur l'incapacité de maintenir l'attention nécessaire à contrôler ses impulsions expliqueraient la plupart des déficits observés dans le groupe des hyperactifs. Il semble également que ce soit sur ces déficits que les stimulants exercent leur effet principal. Des études corrélationnelles et factorielles suggèrent que la même constellation d'habiletés imprègne le comportement des enfants normaux dans plusieurs domaines du fonctionnement cognitif et social.

REFERENCES

BOYDSTUN, J.A., ACKERMAN, P.T., STEVENS, D.A., CLEMENTS, S.D., PETERS, J.E., & DYKMAN, R.A. Physiologic and motor conditioning and generalization in children with minimal brain dysfunction. *Conditional Reflex*, 1968, 3, 81–104.

BURKS, H.F. The hyperkinetic child. *Exceptional Children*, 1960, 27, 18–26.

CAMPBELL, S. Cognitive styles in normal and hyperactive children. Unpublished doctoral dissertation, McGill University, 1969.

CAMPBELL, S.B., & DOUGLAS, V.I. Cognitive styles and responses to the threat of frustration. *Canadian Journal of Behavioural Science*, 1972, 4, 30–42.

CAMPBELL, S.B., DOUGLAS, V.I., & MORGENSTERN, G. Cognitive styles in hyperactive children and the effect of methylphenidate. *Journal of Child Psychology and Psychiatry*, 1971, 12, 55–67.

CARR, L.A., & MOORE, K.E. Norepinephrine release from brain by d-amphetamine *in vivo*. *Science*, 1969, 164, 322–323.

CLEMENTS, S.D., & PETERS, J.E. Minimal brain dysfunctions in the school-age child. *Archives of General Psychiatry*, 1962, 6, 185–197.

Learning, reaction time, activity, and classroom behavior in disturbed children. *American Journal of Orthopsychiatry*, 1970, 40, 615.

SPRAGUE, R.L., & TOPPE, L.K. Relationship between activity level and delay of reinforcement in the retarded. *Journal of Experimental Child Psychology*, 1966, 3, 390–397.

STEIN, L., & WISE, C.D. Release of norepinephrine from hypothalamus and amygdala by rewarding medial forebrain bundle stimulation and amphetamine. *Journal of Comparative and Physiological Psychology*, 1969, 67, 189–198.

STEVENS, D., BOYDSTUN, J., ACKERMAN, P., & DYKMAN, R. Reaction time, impulsivity, and autonomic lability in children with minimal brain dysfunction. *Proceedings of the seventy-sixth annual convention of the American Psychological Association*, 1968, 3, 367–368.

STEVENS, D.A., BOYDSTUN, J.A., DYKMAN, R.A., PETERS, J.E., & SINTON, D.W. Presumed minimal brain dysfunction in children (Relationship to performance on selected behavioral tests). *Archives of General Psychiatry*, 1967, 16, 281–285.

STEWART, M.A. Hyperactive children. *Scientific American*, 1970, 222, 94–98.

STRAUSS, A.A., & LEHTINEN, L. *Psychopathology and education of the brain injured child*. New York: Grune & Stratton, 1947.

SYKES, D.H. Sustained attention in hyperactive children. Unpublished doctoral dissertation, McGill University, 1969.

SYKES, D.H., DOUGLAS, V.I., & MORGENSTERN, G. The effect of methylphenidate (Ritalin) on sustained attention in hyperactive children. *Psychopharmacologia*, 1972, 25, 262–274.

SYKES, D.H., DOUGLAS, V.I., WEISS, G., & MINDE, K. Attention in hyperactive children and the effect of methylphenidate (Ritalin). *Journal of Child Psychology and Child Psychiatry*, 1971, 12, 129–139.

WEISS, G., MINDE, K., WERRY, J.S., DOUGLAS, V.I., & NEMETH, E. Studies on the hyperactive child: A five year follow-up. *Archives of General Psychiatry*, 1971, 24, 409–414.

WERRY, J.S., & SPRAGUE, R.L. Hyperactivity. In C.G. Costello (Ed.), *Symptoms of psychopathology*. New York: Wiley, 1970.

WERRY, J.S., SPRAGUE, R.L., WEISS, G., & MINDE, K. Some clinical and laboratory studies of psychotropic drugs in children – an overview. In W.L. Smith (Ed.), *Symposium on higher cortical functions*. Springfield, Illinois: Thomas, 1969.

WERRY, J., WEISS, G., DOGAN, K., MINDE, K., & DOUGLAS, V.I. Studies on the hyperactive child. VII. Comparison of neurological findings between hyperactive, normal and neurotic children. Paper read at the Canadian Psychiatric Association Annual Meeting, Toronto, June 1969.

WERRY, J., WEISS, G., & DOUGLAS, V.I. Studies on the hyperactive child I: Some preliminary findings. *Canadian Psychiatric Association Journal*, 1964, 9, 120–130.

WERRY, J.S., WEISS, G., DOUGLAS, V.I., & MARTIN, J. Studies on the hyperactive child. III: The effect of chlorpromazine upon behaviour and learning ability. *Journal of American Academy of Child Psychiatry*, 1966, 5, 292–312.

WITKIN, H.A., DYK, R.B., FATERSON, H.F., GOODENOUGH, D.R., & KARP, S.A. *Psychological differentiation*. New York: Wiley, 1962.

First received 15 May 1972

Sex differences in personality structure at age 14*

PHILIP E. VERNON
University of Calgary

ABSTRACT

Though much work has been published on differences between the sexes in Western countries on tests of abilities, interests, and personality traits, little attention has been paid to the organization or structuring of psychological variables among males and females separately. An investigation of nearly 400 Grade 8 students in a Canadian city, including some 10 hours of testing, indicated that the ability factors in the two sexes are closely similar, though their relations to other variables such as age, artistic and scientific interests, social attitudes, and personality tendencies, often differ markedly. For example the psychological significance of field independence and of divergent thinking in the adolescent personality, and the organization of career and other interests, were so different that mixed-sex studies should be discouraged.

Most textbooks, or chapters of books, on individual differences devote considerable space to differences between the sexes in abilities, interests and personality traits (e.g. Anastasi, 1958; Tyler, 1965), and there is no need to recapitulate the well-established facts. Many published tests, at least for adolescents or adults, also provide separate norms for the sexes. Yet it is generally assumed that the main dimensions or factors of ability or personality have the same significance in males and females, and are similarly organized and structured. For example the Primary Mental Abilities batteries, the MMPI, the Study of Values, and Cattell's personality questionnaires are presumed to measure the same abilities or traits in each sex. Counselling or clinical psychologists would, however, be more likely to deny that anxiety, say, or interest in a scientific career, do mean the same thing psychologically in a girl as in a boy, or in a woman as in a man. Likewise, the developmental psychologist tends to find different influences operating in the personality development of boys and girls (cf. Honzik, 1967), and much has been written on the differing identification processes in the acquisition of male and female sex roles (e.g. Kagan, 1964). Rather generally it seems possible to trace fairly straightforward connections between, say, maternal warmth or rejection and the personal characteristics of boys; whereas in girls, personal behaviour depends more on the immediate social context (cf. Schaefer & Bayley, 1963). Different personality patterns, again, are associated with verbal, number, and spatial abilities in boys and girls, according to Ferguson and Maccoby (1966).

* The writer is greatly indebted to The Canada Council, Social Sciences Division, for financial support of this study. A brief version of this paper was presented at the Canadian Psychological Association meeting in Montreal, June 1972.

Impulsivity - reflectivity and differential reinforcement of low rates (DRL) performance

Norman Stein and Richard Landis

Syracuse University

Paper presented at the meeting of the Eastern Psychological Association, Washington, D.C., 1973.

The present study is derived from two, largely independent, areas of research; first, the cognitive style literature on impulsivity--reflectivity and, second, the operant conditioning literature on differential reinforcement of low rates (DRL) schedules.

Regarding the area of cognitive style, Kagan (1965) has isolated a conceptual tempo dimension called "impulsivity--reflectivity" which underlies modes of problem solving in children. The task used to differentiate these groups is called the Matching Familiar Figures (MFF) Test (See Figure 1). On this test the child points out the one stimulus, from among six highly similar variants, that is identical to a standard. Subjects scoring above the sample median for number of errors and below the median for decision time are designated as impulsive, while subjects scoring above the median for decision time and below the median for number of errors are designated as reflective. Research has consistently shown that impulsive children perform relatively poorer across a variety of motor, perceptual, and conceptual tasks which typically require an inhibition of the choice response in order to optimize performance. If this is the case, then the conceptual tempo dimension should be predictive of differences in performance on a learning task which requires temporal spacing of responses in order to maximize reinforcement.

One such task is derived from operant conditioning methodology and consists of responding on a telegraph key programmed to provide reinforcement according to a DRL 6 sec. schedule. The schedule provides reinforcement contingent upon responses which terminate inter-response times (IRTs) of 6 sec. or longer. Responses terminating IRTs shorter than 6 sec. are not reinforced and, in addition, reset the timer governing reinforcement.

Thus, in order to maximize reinforcement, a child must refrain from engaging in the reinforced response for at least 6 sec. The alternate behavior which occurs during the IRT is called collateral behavior. In the present experiment, collateral behavior is measured by presses on either of three telegraph keys which are adjacent to the directly reinforced response - key. Previous research has shown that under such an arrangement, accomodation to the DRL schedule requirement is accompanied by regular and frequent occurrences of unscheduled collateral behavior and that this collateral behavior, through the effects of adventitious reinforcement, can serve a mediating function that regulates the temporally discriminated performance (Stein and Landis, 1973).

The present study documents the performance of cognitively impulsive and reflective children operating under a DRL schedule of reinforcement by measuring response rates, reinforcement rates, reinforcements per response rate (i.e., efficiency), distribution of IRTs, and the behavioral mechanisms used to mediate successively reinforced responses. In addition, it sought to examine the role of instructions about the requirements for reinforcement on these variables.

METHOD

Subjects: The MFF was administered to 73 third-grade boys of a suburban elementary school. From this pool, 20 impulsive and 20 reflective children were identified. The mean age was 8.94 years and the mean IQ was 107.6. There were no significant differences between reflective and impulsive Ss on these characteristics.

Apparatus. The Ss were seated before a chassis containing four adjacent telegraph keys mounted 9.5 cm apart. Each key required a downward force of 438 gm to be electronically recorded as a response. A four-digit add counter was mounted in the middle of the chassis to indicate the cumulative number of reinforcements (points). Points were backed up by M & M candy reinforcers. White noise (60 db) was continuously presented through headphones worn by the S and a sound attenuating chamber housing the programming equipment acoustically isolated the S from the equipment. The third key from the Ss' left (Key 3) was programmed to produce reinforcement according to a DRL 6-sec. schedule.

Procedure. The general design involved four equal sized groups: two groups of impulsive Ss, and two groups of reflective Ss, one group from each received instructions about the DRL requirements and one group received no instructions about these requirements. Initially, to all Ss, the task was introduced as a game in which they should try to cumulate points and that each point equalled one M & M. Each point was also to be accompanied by a flash of a red light above the counter. Next, to assess S's baseline rate of key pressing all Ss were asked to practice using the keys but told they would not be able to obtain points at this time. After a two minute baseline phase, the two Instruction Groups were explicitly told the nature of the contingencies that would be in effect during the next 30 minutes. Essentially, they were told "...if you press the second key from your right, wait for 6 seconds before you press it again, then you'll get a point every time. If you press it before 6 seconds are up, then you will not get a point and will have to wait another 6 seconds before you can get a point."

RESULTS

Figure 2 shows the mean number of responses on Key 3, the key programmed on the DRL 6 sec. schedule, during baseline and during successive 5 minute intervals throughout training. No difference in response rate between the groups occurred during baseline ($F = 0.31$, $df = 3/36$, $p > .10$). However, throughout DRL training significant main effects were obtained for Instructions ($F = 6.65$, $df = 1/36$, $p = \leq .01$), Conceptual Tempo ($F = 4.51$, $df = 1/36$, $p = \leq .04$), and Time ($F = 3.19$, $df = 5/180$, $p = \leq .01$). Comparison of group means by t-tests showed that instructions about the contingencies significantly reduced the response rate of impulsives (Group II) compared to uninstructed impulsives (Group INI) ($p \leq .005$). The Uninstructed Reflectives (Group RNI) did not differ significantly from the Instructed Reflectives (Group RI). Though both Instructed groups responded with similar rates, the difference between INI and RNI groups was significant ($p \leq .01$). It was also found that while the rate of responding decreased from the first to the last 5 minutes of training for both the INI ($t = 2.73$, $df = 180$, $p = \leq .01$) and RNI ($t = 2.50$, $df = 180$, $p = \leq .05$) groups, both instructed groups responded at essentially the same low rate throughout DRL training.

Figure 3 shows the mean number of collateral responses per minute on the three nonreinforced keys during baseline and during successive five minute blocks throughout DRL training. Again, no significant difference between groups was obtained during the baseline period ($F = 0.59$, $df = 3/36$, $p > .10$). During DRL training, there was a significant main effect for Instructions ($F = 13.04$, $df = 1/36$, $p = \leq .001$), in which the RNI Group responded with reliably higher rates than the II Group ($p \leq .005$). Similarly, the RNI Group emitted significantly more collateral responses than the RI group ($p \leq .05$). Significant interactions were also obtained for both Time X Instructions ($F = 4.14$, $df = 5/180$, $p < .001$) and Time X Conceptual Tempo ($F = 2.66$, $df = 5/180$, $p < .02$). Analyses of simple main effects revealed that while the two instructed groups responded with significantly lower rates than the two uninstructed groups during each time period ($p < .05$), only the uninstructed groups showed significant changes over time ($F = 16.36$, $df = 5/180$, $p < .01$). Thus, together with the findings on Key 3 responding, it was found that only the uninstructed groups exhibited changes in the rate of responding during exposure to the DRL contingencies. However, while the collateral response rates significantly increased from the beginning to the end of training for both the INI ($t = 4.73$, $df = 180$, $p < .001$) and the RNI ($t = 7.36$, $df = 180$, $p < .001$) groups, their rate of Key 3 responding had significantly decreased during the same time periods.

The direct comparisons between the rate of responding on Key 3 and on the collateral keys (1, 2, 4) for all groups are seen clearly in Figure 4. Series of one way ANOVAs indicated no differences within the INI, II, and RI groups in the response rates across the keys. However, the RNI Group tended to respond less frequently on Key 3 than on the collateral keys ($F = 2.68$, $df = 3/36$, $p < .10$) and at a rate comparable to that of the instructed groups.

The proportion of IRTs $\geq$ 6 sec for Key 3 responses, which reflects the efficiency of performance — that is, the percent of Key 3 responses that were reinforced, is shown in Figure 5. The same essential and almost identical relationships, were obtained for the sheer number of reinforcements per minute. Again, the four groups were found to be comparable in the proportion of IRTs $\geq$ 6 sec emitted during baseline ($F = 1.65$, $df = 3/36$, $p > .10$). A significant main effect was obtained for Instructions for both the efficiency

($F = 17.44$, $df = 1/36$, $p = < .001$) and the reinforcements per min ($F = 10.19$, $df = 1/36$, $p = < .003$) measures. Instructions significantly increased the efficiency ($p < .02$) and the number of reinforcements obtained ($p < .005$) by Impulsives. In contrast, Instructions enhanced the efficiency by Reflectives ($p < .03$) but did not correspondingly increase the rate of reinforcement ($p > .15$). The Time x Instructions interaction was also significant for the efficiency ($F = 3.77$, $df = 5/180$, $p = .003$) and reinforcements per min ($F = 4.21$, $df = 5/180$, $p = .001$) measures. Comparisons of the first and last time blocks revealed that both the INI and RNI Groups became more efficient and earned reinforcement at a faster rate as a result of extended exposure to the DRL contingency ($p \leq .001$), the RI Group also improved significantly on these measures ($p < .01$), while the II Group responded no more efficiently and earned no more reinforcements at the end than at the beginning of DRL training.

Although the INI group exhibited a significantly higher Key 3 response rate than the RNI group, the main effects for conceptual tempo on the efficiency ($F = 3.04$, $df = 1/36$, $p = .08$) and reinforcement rate ($F = 3.03$, $df = 1/36$, $p = .08$) measures failed to attain traditionally acceptable level of significance. Nevertheless, further analyses revealed that while the Instructed groups did not differ, the RNI Group responded more efficiently ($p < .05$) and obtained more reinforcements ($p \leq .025$) than the INI Group.

To assess the structure of the temporal performance for each group, the relative frequencies of IRT and the IRT/OP on Key 3 were obtained for the baseline period and for the first, middle, and last five min blocks of training. The results are shown in Figure 6. As shown, during Baseline, Impulsives emitted relatively fewer long IRTs than Reflectives and, given the opportunity, had a high probability of terminating IRTs of about 4 sec. The IRT/OP distributions for the Reflectives during Baseline were flat, indicating that the probability of emitting an IRT of any duration between 0-16 sec was the same. Most noticeable was the immediate elimination of short IRTs by the introduction of instructions and the rapid development of an accurate temporal discrimination by the Instructed Reflective and Impulsive Groups. The IRT distributions of the II group remained essentially unchanged throughout training while the RI Group showed a more gradual development of a precise discrimination. In contrast, the Uninstructed Impulsive and Reflective Groups generated considerably more IRTs in the shorter categories throughout training. In neither group was bursting (i.e., 0-2 sec IRTs) eliminated entirely; however, for Group INI, approximately 30% of all IRTs were in the 0-2 sec category during the last five min of training. In contrast, this percentage is equalled by the RNI Group only during the first five min of training. Thereafter, the percentage of 0-2 sec IRTs decreased steadily until, at the end of training, only 10% of the IRTs were less than 2 sec.

The proportion of bursts (i.e., a sequence of two or more responses with IRT ≤ 2 sec.) immediately following reinforced responses was determined. It was found that, on the average, for Group INI 31% of the reinforced responses were followed immediately by bursting while for Groups II, RNI, and RI the mean proportions were 2%, 6% and 0%, respectively. Statistical analysis indicated no differences between these latter groups but each was significantly different from Group INI.

DISCUSSION

Since there were no differences between reflective and impulsive Ss on the baseline measures of motor activity, the differences in responding between these groups during DRL must be accounted for in terms other than simply differential predispositions to respond at faster or slower rates. The present results suggest that whether or not cognitively impulsive and reflective children exhibit high or low response rates depends upon the "stimulus structure" or lack of ambiguity present for the child in the learning situation. When the DRL task was structured and made unambiguous through explicit instructions about the requirements for reinforcement, there were no reliable differences between impulsives and reflectives. However, when the DRL task was unstructured and ambiguous, such that the relationships between the task, the S's behavior, and the consequences of the behavior were unspecified, reliable differences between reflective and impulsive children emerged. Under these latter conditions, impulsive children emitted higher rates of the reinforced response, tended to obtain fewer reinforcements, were less efficient, and responded with equal frequencies on all the response keys. In contrast, uninstructed reflectives emitted a significantly lower response rate on the reinforced key in comparison to the collateral keys.

The higher response rate exhibited by the uninstructed impulsives was not a function of a higher general arousal or activity level. Rather, the higher rates were the result of rapid repetitions of the reinforced response (i.e., response bursts) and their occurrence was specifically related to the effects of reinforcement. Rapid bursts of responding typically occurred under two conditions: (1) for both uninstructed reflective and impulsive children bursting

R as S^D for responding

occurred following nonreinforced Key 3 responses, and (2) for uninstructed impulsive children only, bursting also occurred following reinforced Key 3 responses. In the first instance, bursting probably reflected an accommodation to an unreceived but anticipated reward. In the second instance, which was specific to uninstructed impulsive Ss, bursting appeared as an operant under the discriminative control of the prior reinforcement. Thus, for uninstructed impulsive Ss, reinforcement may have served as an S_D for rapid repetition of the same response. In contrast, with uninstructed reflectives, reinforcement may have served as an S_D to engage in collateral behavior which effectively mediated the IRT.

The observation that reinforcement may control different behavioral tendencies for reflective and impulsive children in previously unstructured situations is also relevant to the proposal that impulsives fail to discriminate as accurately those situations in which it is appropriate to respond slowly. The data shown in Fig. 4 support this hypothesis, such that the INI Group responded with equal frequency on all keys, while the RNI Group responded less frequently on the reinforced key than on the collateral keys. Thus, impulsives exhibited a flat or broad gradient of generalization and reflectives exhibited a sharp gradient of generalization with a peak at the key directly associated with reinforcement. Since a steep generalization gradient reveals the presence of a discrimination, it is possible that reflectives discriminated stimuli associated with reinforcement from those associated with non-reinforcement. The broad generalization gradient obtained from the uninstructed impulsive Ss reveals the absence of such a discrimination. This interpretation, however, is limited by the finding that impulsives obtained fewer reinforcements than

FIG. 2

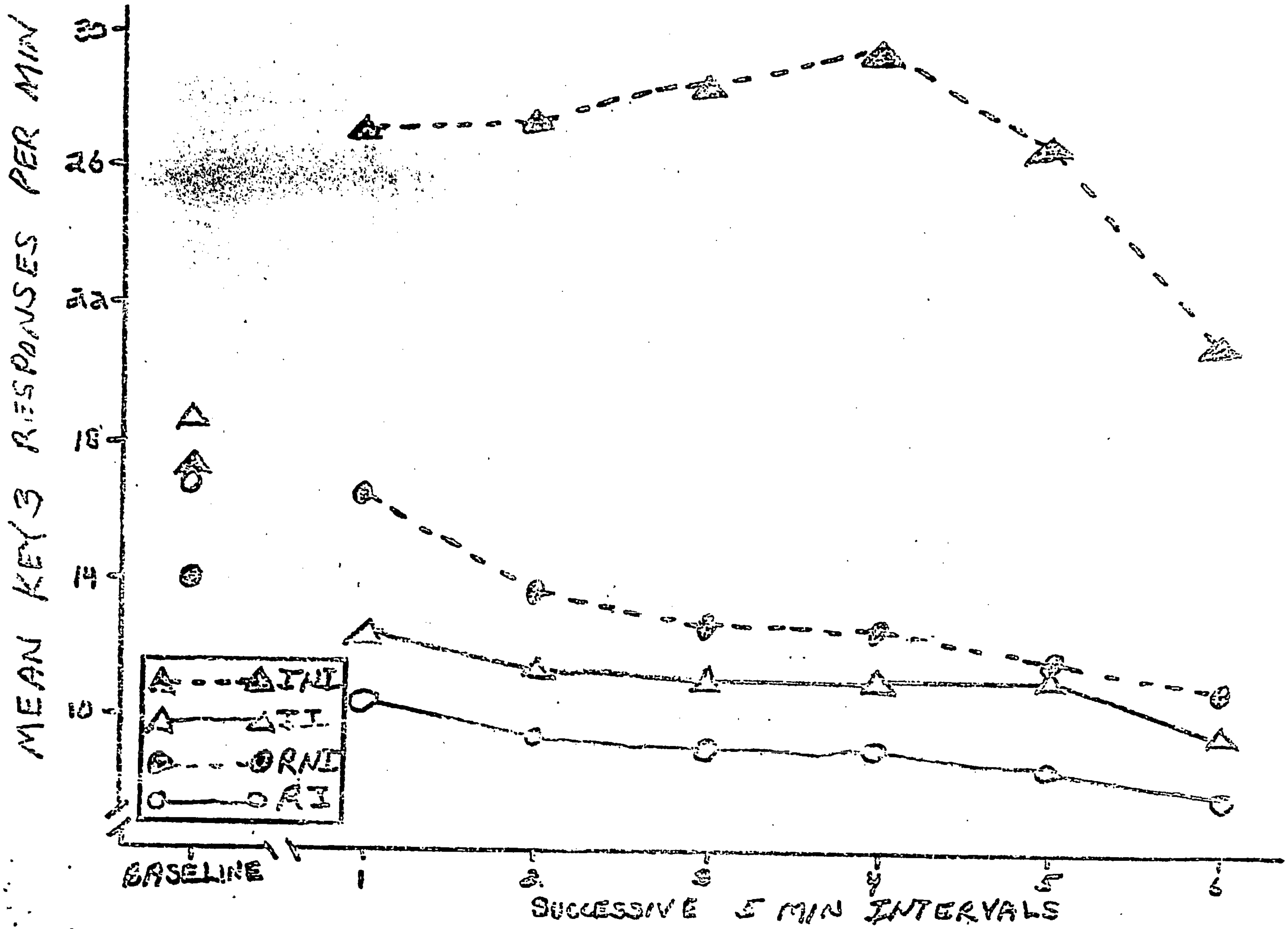

MEAN KEY 3 RESPONSES PER MIN
30
26
22
18
14
10
INI
II
RNI
RI
BASELINE
1
2
3
4
5
6
SUCCESSIVE 5 MIN INTERVALS

FIG. 3

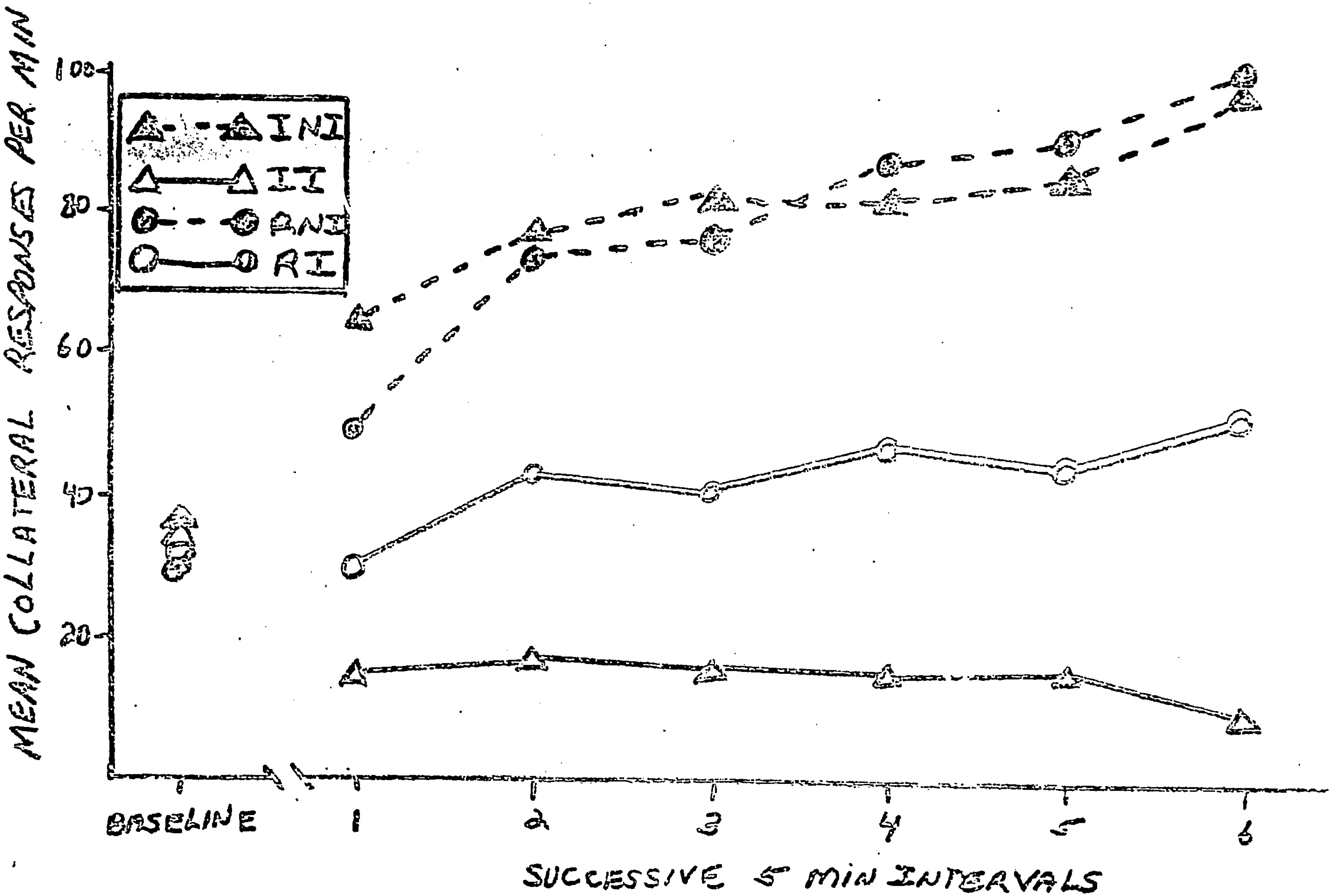

MEAN COLLATERAL RESPONSES PER MIN
100
80
60
40
20
INI
II
ANI
RI
BASELINE
1
2
3
4
5
6
SUCCESSIVE 5 MIN INTERVALS

FIG. 4

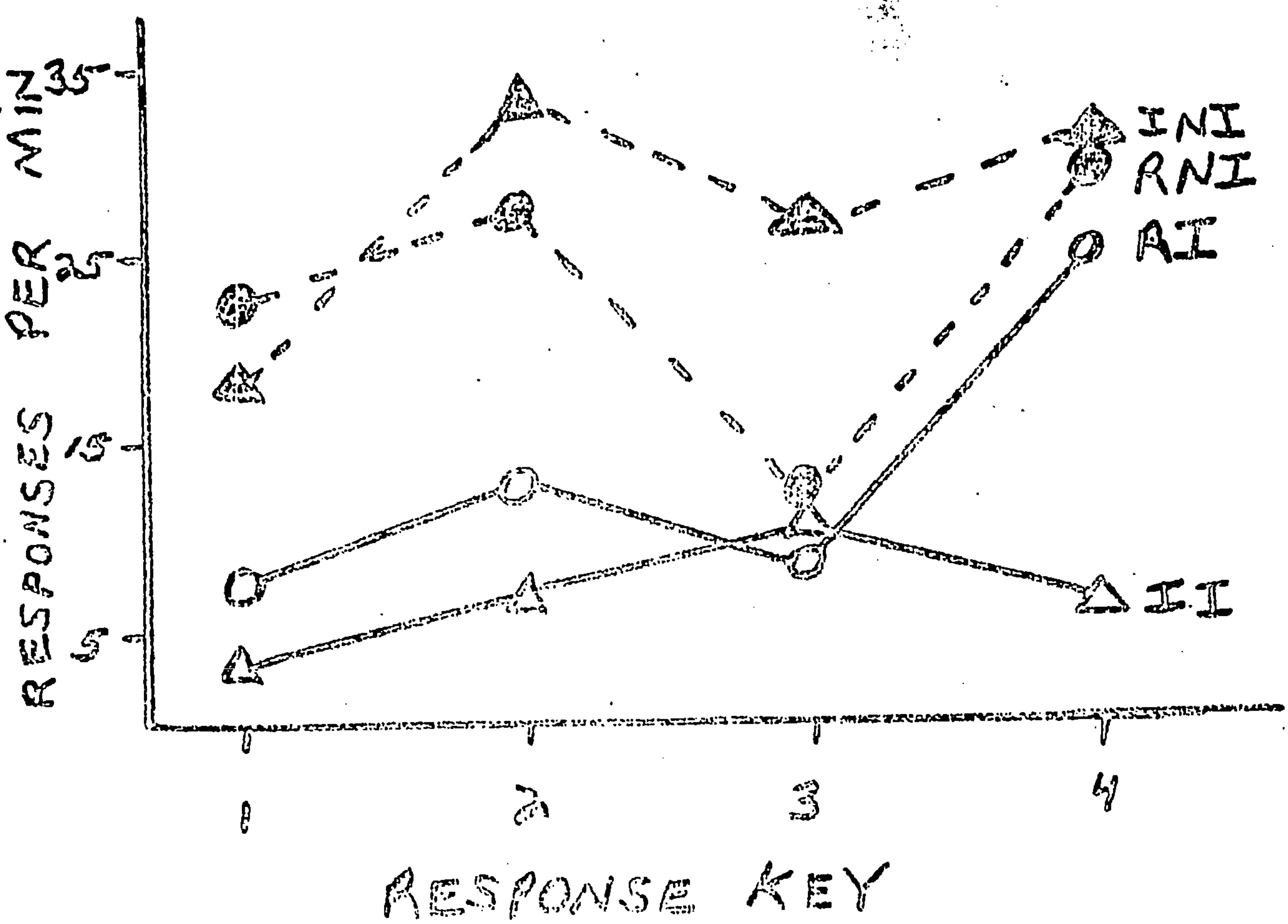
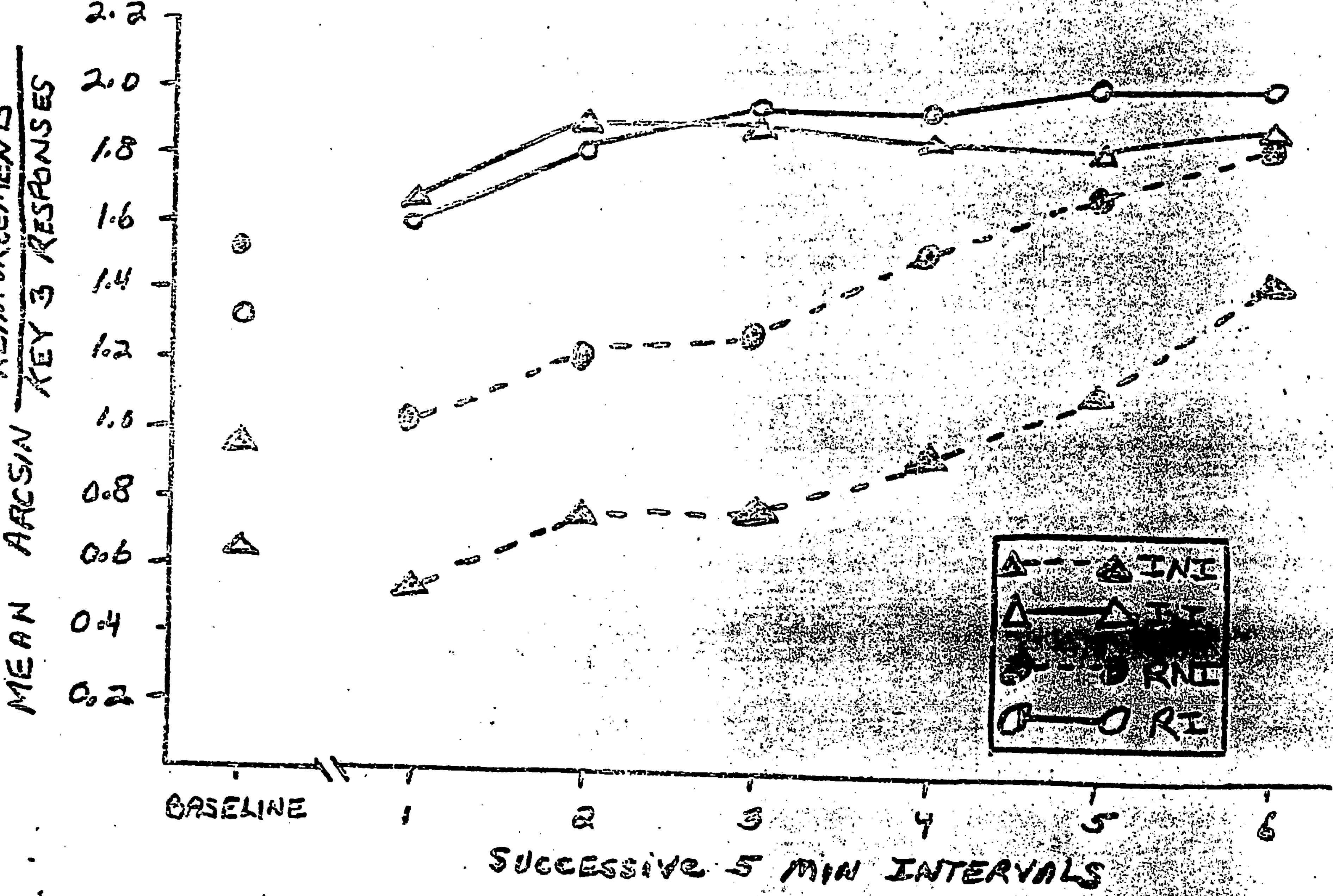


FIG. 5


MEAN ARCSIN REINFORCEMENTS / KEY 3 RESPONSES
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
BASELINE
1
2
3
4
5
6
SUCCESSIVE 5 MIN INTERVALS
INI
RNI
RI

reflectives. Thus, it is unclear whether impulsives obtained fewer reinforcements because they may fail to discriminate between stimuli associated with reinforcement from those which are not, or fail to discriminate because they receive fewer reinforcements.

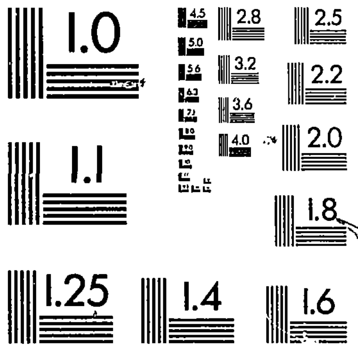
REFERENCES

Kagan, J. Impulsive and reflective children: Significance of conceptual tempo. In J. Krumboltz (Ed.) Learning and the educational process. Chicago: Rand McNally, 1965.

Stein, N. and Landis, R. Mediating role of human collateral behavior during a spaced-responding schedule of reinforcement. Journal of experimental psychology, 1973, 97 (1), 28-33.

Qu K854903 84 2 11

IMAGE EVALUATION TEST TARGET (MT-3)



150mm

6″

PHOTOGRAPHIC SCIENCES CORPORATION

770 BASKET ROAD
P.O. BOX 338
WEBSTER, NEW YORK 14580
(716) 265-1600

K854903

See page 131

J. Child Psychol. Psychiat., Vol. 12, 1971, pp. 129 to 139. Pergamon Press. Printed in Great Britain. 851

ATTENTION IN HYPERACTIVE CHILDREN AND THE EFFECT OF METHYLPHENIDATE (RITALIN)*

DONALD H. SYKES,† VIRGINIA I. DOUGLAS, GABRIELLE WEISS and KLAUS K. MINDE

Departments of Psychology and Psychiatry, McGill University and the Montreal Children's Hospital, Montreal, Canada

CHILDREN are frequently described as overactive, particularly by parents and teachers. Lapouse and Monk (1958) estimated that some 50 per cent of all children between the ages of 6 and 12 yr are described by their mothers as highly active although their activity is not necessarily a problem. Stewart *et al.* (1966) have estimated that 4 per cent of all school-age children show a level of activity sufficiently excessive and sustained as to be a serious source of complaint at both home and school. Such children have been labelled "hyperactive" by clinicians.

Given the problems surrounding the definition and measurement of activity level, particularly in humans (Cromwell, Baumeister and Hawkins, 1963), it is by no means surprising that there is little agreement on the definition of hyperactivity. For some investigators it implies a greater quantity of movement (Schulman, Kaspar and Throne, 1965), but this has been questioned by others (Werry and Sprague, 1970). Cromwell *et al.* (1963) have suggested that the "overactivity" of "hyperactive" children may be a reflection of the short attention span and rapidly changing goal directions of such children. Thus, these investigators argue that hyperactive children may be thought of as children whose behaviour is fragmented or disorganized and continually changing direction such that an impression of a high level of activity is created.

Certainly, the inability of hyperactive children to maintain their attention toa task has frequently been remarked in the clinical literature. However, other than the clinical statements and the remarks of teachers and parents, there are few experimental studies which clearly show an attentional deficit in such children. In fact, there is no reliable evidence that they are impaired on the usual tests of attention, such as the coding subtest of the Wechsler Intelligence Scale for Children (Douglas, Weiss and Minde, 1969). It has been suggested with respect to this latter test, that if the attentional impairment took the form of brief lapses in attention, then performance on the coding test need not necessarily be impaired. The nature of the

*The research was supported by a research grant from the Ciba Pharmaceutical Company, Dorval, Quebec, to Dr. G. Weiss.

†Present address: Department of Psychology, The Queen's University of Belfast, Belfast, Northern Ireland.

Accepted manuscript received 16 *April* 1971

coding test is such as to allow the subject to compensate for any brief lapses by working rapidly in between lapses (Rosvold *et al.*, 1956). In contrast, impaired performance due to momentary lapses in attention should be noted on a task in which the arrival of significant information is unpredictable. A task designed to meet these requirements was constructed by Rosvold *et al.* (1956), namely the Continuous Performance Test.

This task requires the subject to monitor a screen on which letters appear at regular intervals and to make a response whenever a previously specified stimulus appears. The test is similar to many of the tasks used in vigilance studies. Any momentary lapse in attention which is coincident with the appearance of a significant stimulus would appear as an error of omission. Thus, the primary purpose of the present study was to examine the performance of a group of hyperactive children relative to the performance of a group of normal control children on a task susceptible to momentary lapses in attention.

It has been observed that motor restlessness in adults increases with time on a vigilance task (Baker, 1960), although no relationship was found between actual performance on the task and the increase in restlessness. In the present study the motor restlessness of hyperactive and normal children while seated during the attention task was also examined, using a stabilimetric cushion after the design of Sprague and Toppe (1966).

The third purpose of the study was to investigate the effect of methylphenidate (Ritalin), a central nervous system stimulant, on attention in hyperactive children. A number of investigators have found this drug particularly efficacious in the treatment of hyperactive children (Campbell, Douglas and Morgenstern, 1969; Cohen, Douglas and Morgernstern, 1969; Conners, Eisenberg and Sharpe, 1964; Douglas *et al.*, 1969; Eisenberg, Conners and Sharpe, 1965; Knights and Hinton, 1969; Sprague, Barnes and Werry, 1970; Weiss *et al.*, 1970). However, there is still some question as to its specific effects on sustained attention in hyperactive children (Conners *et al.*, 1964).

METHOD

Subjects

Since the reliable and valid measurement of activity has proved to be a difficult task (Cromwell, Baumeister and Hawkins, 1963), no exact definition of "hyperactive child" in terms of a measured quantity of movement or activity has yet been made. Consequently, the definition of such children in the present study was an operational one based on the selection criteria used. For a child to be included in the study, hyperactivity had to be the major complaint. That is, both the child's parents and teacher had to specify the overactivity of the child as their major complaint and the reason for the referral. Furthermore, such hyperactivity had to have been present as a chronic problem since early childhood and be sustained throughout the day. To ensure as homogeneous a group as possible, children diagnosed by the child psychiatrists as psychotic, epileptic or brain-damaged, or whose major-presenting symptom was behaviour disturbance of an emotional nature, were excluded from the study.

Forty children (34 males and 6 females) who met the above criteria were included in the study. The ages of the *S*s ranged from 5 to 12 yr (mean age = 8 yr; *s.d.* = 1 yr 9 months) and their intelligence quotients (as measured by the Wechsler Intelligence Scale for Children (Wechsler, 1949)) were not less than 80 (mean *I.Q.* = 111; *s.d.* = 11).

A control group of 19 normal children, matched for age, sex and *I.Q.* with 19 of the hyperactive children, was also tested. The two groups did not differ on mean age ($t = 0{\cdot}003$, $df = 18$, *NS*) or *I.Q.* ($t = 0{\cdot}018$, $df = 18$, *NS*). The control children were selected from the normal school population of the City of Montreal.

All of the children were English speaking, living at home with at least one parent and were attending regular school classes.

Measures

Attention task. The Continuous Performance Test (C.P.T.) is an experimenter-paced task, that is *E* controls the arrival and duration of the task stimuli. A series of letters are presented one at a time on a screen. *S* is required to monitor the screen and respond whenever a specified stimulus, the significant stimulus, appears. Three variables were manipulated in the present study, namely task stimuli, distraction, and interstimulus interval.

There were three separate series of task stimuli, referred to as the *X* the *AX* and the Form sequences. These task stimuli differed in the following respects. In the *X* sequence 12 letters appeared on the screen in a random serial order, the letter *X* being the significant stimulus to which *S* had to respond. In the *AX* sequence the same 12 letters as in the *X* sequence were used, but here the significant stimulus to which *S* responded was the letter *X* only when it was immediately preceded by the letter *A*. The stimuli for the Form sequence consisted of seven geometric shapes in one of five colours. The significant stimulus to which *S* was required to respond was a red triangle.

Each of the three sequences (*X*, *AX* and Form) were presented under two conditions of distraction, minimal and intermittent noise. In the minimal condition *S*'s room was kept as free as possible of extraneous distracting noise. In the intermittent condition white noise (80 decibels) was piped into *S*'s room at frequent random intervals.

Each of the three sequences was displayed at two interstimulus intervals, 1·0 and 1·5 sec. Thus, for both groups of *S*s there were three sequences (*X*, *AX* and Form), two distraction conditions (minimal and intermittent noise) and two interstimulus intervals (1·0 and 1·5 sec), making 12 trials in all for each *S* (i.e. a four-way complete factorial design).

On each of the 12 trials a total of 200 stimuli were presented. Thus, at the 1·0 sec interstimulus interval a trial was continuous for 3·3 min and at the 1·5 sec interstimulus interval for 5 min.

Motor restlessness. Restlessness while seated was measured with a stabilimetric cushion (Sprague *et al.*, 1966). Movements of the *S* on the seat in the left–right and front–back directions activated microswitches placed underneath the cushion. The microswitches were connected to digital counters which gave a total score based on the movements of the *S* in all directions.

Procedure

All *S*s were tested individually. Each *S* visited the hospital on two occasions to complete the initial (pre-drug) testing. During the first session 6 of the 12 trials were given, the remaining 6 trials being given during the second session. The trials were randomized such that order effects were controlled. During the actual testing *S* was alone in the test room, while *E* was in an adjoining room monitoring the equipment and *S*'s behaviour through a one-way screen.

Upon completion of the initial testing each of the 40 hyperactive children was assigned, using a randomized code provided by the Ciba Pharmaceutical Company, to one of two groups, Active Drug or Placebo, using the double-blind technique. The two psychiatrists then titrated the drug for each of their patients until the optimum clinical effect, based on the parents' report of the child's behaviour and the psychiatrist's evaluation of that report, was reached. For most of the children, this optimum clinical effect was reached when the dosage was in the region of 30–40 mg per day. Each hyperactive child returned for retesting (again for two sessions) while on either the active drug or a placebo, 5–7 weeks after completion of the initial testing. The code was not broken until on-drug testing was completed. The control *S*s were seen only for the initial testing (i.e. for only two of the four sessions).

RESULTS

On the C.P.T. there were two dependent measures, the absolute score and the error score. The absolute score is a measure of the *S*s accuracy in detecting the significant stimuli (i.e. No. correct responses/Total no. significant stimuli presented × 100) and the error score is the number of responses to non-significant stimuli.

Hyperactive-control comparisons

Two separate four-way analyses of variance, repeated measures design (Winer, 1962) were used to compare the absolute scores (Table 1a) and the error scores (Table 1b) of the hyperactive and normal children. The four factors were, (1) Groups—hyperactive vs. control, (2) Sequence—X, *AX* and Form, (3) Interstimulus interval—1·0 sec vs. 1·5 sec, and (4) Distraction—noise vs. no noise.

A significant Groups factor ($p < 0{\cdot}025$) was obtained when the absolute but not the error scores were analysed. This finding indicated that the hyperactive children made significantly fewer correct responses than did the control children (mean hyperactive group = 75·54 per cent, mean control group = 84·49 per cent). However, the groups did not differ with respect to overall number of errors (mean hyperactive group = 6·69, mean control group = 6·12).

Two other main factors were also found to be significant, namely Interstimulus Interval and Sequence. The significant interval factor indicated that both groups of children (hyperactive and control) made more correct responses and fewer errors when the slow (1·5 sec) rather than the fast (1·0 sec) interstimulus interval was used (1·5 sec interval: mean correct both groups = 85 per cent, mean errors = 5·57: 1·0 sec interval: mean correct = 74·5 per cent, mean errors = 7·73).

A main effect for Sequence was found only when absolute scores were analysed and indicated that both groups of children made more correct responses for the X

TABLE 1. TWO FOUR-WAY ANALYSES OF VARIANCE (REPEATED MEASURES) OF THE ABSOLUTE AND ERROR SCORES OBTAINED BY HYPERACTIVE AND NORMAL CHILDREN ON THE CONTINUOUS PERFORMANCE TEST

Source	*df*	(1a) Absolute score *MS*	*F*	*p*	(1b) Error score *MS*	*F*	*p*
Between *S*s							
Groups	1	9126·32	6·76	< 0·025	37·06	0·25	
S (within groups)	36	1350·33			147·80		
Within *S*s							
Sequence	2	903·37	3·64	< 0·05	32·00	1·21	
Groups × sequence	2	363·14			25·04	0·95	
Sequence *S* (within groups)	72	247·95			26·35		
Distraction	1	1·98			8·98	1·30	
Groups × distraction	1	3·51			0·71	0·10	
Distraction *S* (within groups)	36	35·66			6·93		
Interstimulus interval	1	13882·15	158·14	< 0·005	313·34	20·39	< 0·005
Groups × interval	1	0·00			130·56	8·49	< 0·01
Interval *S* (within groups)	36	87·78			15·37		
Sequence × distraction	2	6·64			10·86	1·55	
Groups × sequence × distraction	2	22·40			14·42	2·06	
Sequence × distraction *S* (within groups)	72	33·44			7·01		
Sequence × interval	2	318·03	3·96	< 0·025	15·13	0·92	
Groups × sequence × interval	2	97·65			1·23	0·08	
Sequence × interval *S* (within groups)	72	80·22			16·37		
Distraction × interval	1	94·87			0·43	0·04	
Groups × distraction × interval	1	72·64			0·00	0·00	
Distraction × interval *S* (within groups)	36	49·88			10·25		
Sequence × distraction × interval	2	27·52			1·49	0·19	
Groups × sequence × distraction × interval	2	43·24			0·60	0·08	
Sequence × distraction × interval *S* (within groups)	72	54·83			7·72		

(mean = 82·29 per cent) than for the Form (mean = 80·31 per cent) and *AX* (mean = 77·44 per cent) sequences respectively. Mean errors did not differ significantly across the three sequences (mean *X* sequence = 6·74, mean Form sequence = 5·88, mean *AX* sequence = 6·59).

The main effect for Distraction did not reach significance when either absolute or error scores were anlysed. Thus it may be concluded that the distraction factor did not produce any significant decrement in the C.P.T. performance of either group.

Only two interactions reached significance, Groups by Interstimulus Interval (for error scores only) and Sequence by Interstimulus Interval (absolute scores only). The Groups by Interval Interaction is represented in Fig. 1, and indicates that although the hyperactive and control children made a similar number of errors at the fast interval (1·0 sec), the control children made significantly fewer errors at the slower interval (1·5 sec) than the hyperactive children.

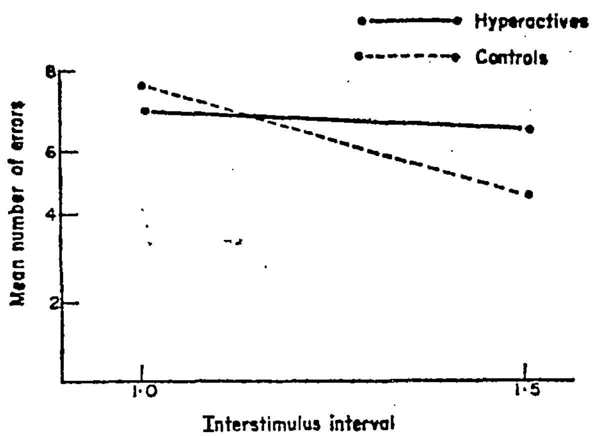


FIG. 1. Incidence of errors on the continuous performance test of the hyperactive and control groups at two interstimulus intervals.

The significant Sequence by Interval interaction indicated that the difficulty of the three sequences, *X*, *AX*, and Form depended on the interstimulus interval used. For the 1·0 sec interval more correct responses were made by both groups to the *X* (mean = 78·36) than to the Form (mean = 74·46) or *AX* (mean = 70·67) sequences. When the 1·5 sec interval was used the order was Form (mean = 86·16), *AX* (84·21) and *X* (82·22).

Motor restlessness

The total activity scores for each *S* accumulated during each of the two testing sessions were first transformed to logarithms to reduce the large variation between *S*s (Winer, 1962) and then analysed by means of a two-way analysis of variance, repeated measures design. The two factors were Groups (hyperactive vs. control) and Session (1st vs. 2nd visit).

The hyperactive children accumulated significantly greater activity scores than the control children (F for groups = 6·84, $df = 1/36$, $p < 0·05$). The main effect for sessions also reached significance, indicating that for both groups restlessness increased during the second session ($F = 192·27$, $df = 1/36$, $p < 0·001$). Moreover, the significant Groups by Session interaction ($F = 7·81$, $df = 1/36$, $p < 0·01$) indicated that the restlessness of the hyperactive children increased at a faster rate from the 1st to the 2nd session than did that of the control children (Fig. 2).

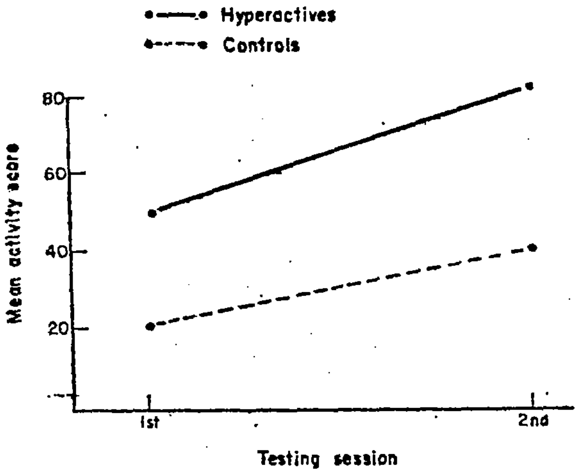


FIG. 2. Restlessness scores of hyperactive and control children during two testing sessions.

Drug comparisons

Despite the fact that the hyperactive children were randomly assigned › either the Active Drug or Placebo group ($N = 20$ per group) and a double-blind technique was used, it was found that those hyperactive children receiving the active drug had actually obtained poorer pre-drug C.P.T. scores than had those placed on a placebo. Consequently, it was necessary to use analyses of covariance (covarying for initial C.P.T. performance) on both absolute and error scores when evaluating the effect of methylphenidate compared to placebo.

Two-way analyses of covariance (Winer, 1962) were completed on each of the three task sequences (X, AX, and Form) separately. The two factors were Groups (drug vs. placebo) and Interval (1·0 sec vs. 1·5 sec). Distraction (noise vs. no noise) was not used as a main factor in these analyses since the previous analyses (Table 1A and 1B indicated that it did not affect performance.

For all three task sequences, the active drug group had a significantly higher absolute score than did the placebo group ($p < 0·005$ in all three cases). The active drug group also made significantly fewer errors than the placebo group on the X ($p < 0·005$) and Form ($p < 0·05$) but not the AX sequences.

The main effect for Interval was significant for all three task sequences when absolute scores were analysed. As in the initial analyses, this finding indicated that

*S*s made more correct responses when the slow rather than the fast interstimulus interval was used. They also made fewer errors when the slow interval was used, this effect reaching significance only in the analysis of error scores on the *AX* task.

The Groups by Interval interaction was not significant in the analyses of absolute scores for any of the three tasks. When error scores were analysed, however, this interaction reached significance in the *X* condition ($p < 0{\cdot}5$) and showed a trend in the *AX* condition ($p < 0{\cdot}10$). This interaction is illustrated in Fig. 3 and indicates that whereas the active drug group made fewer errors at the slow than at the fast interstimulus interval, the placebo group made a similar number of errors at both speeds. The behaviour of the active drug group in this respect is reminiscent of the behaviour of the normal control children at the slower speed (see Fig. 1).

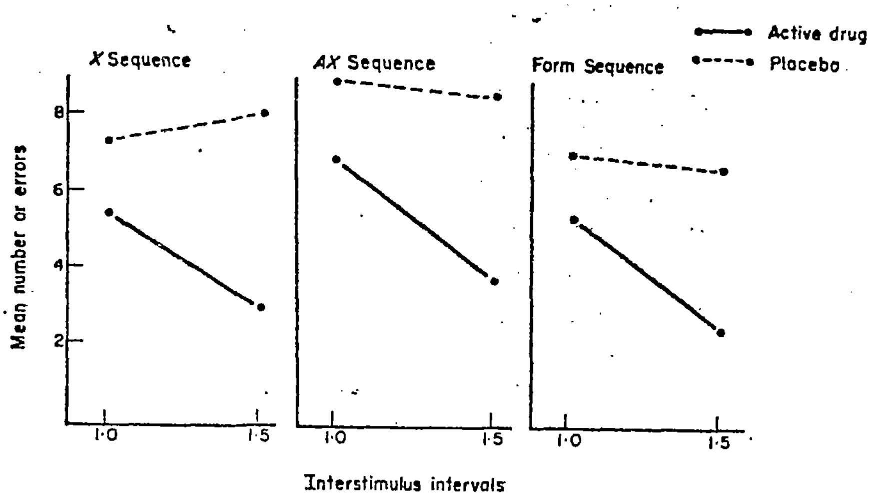


FIG. 3. Incidence of errors on the continuous performance test of the active drug and placebo groups at two interstimulus intervals.

DISCUSSION

The results of the present study lend support to the hypotheses that hyperactive children are deficient with respect to their ability to maintain attention to a task and are physically more restless than normal children.

On a three-part task (*X*, *AX* and Form) requiring the monitoring of a screen, the hyperactive children detected significantly fewer of the significant stimuli. They also made significantly more responses to non-significant stimuli than did the control children at the slow, but not at the fast interstimulus interval. It should be noted, however, that the performance of the hyperactive children was not as impaired as that of the brain-damaged and centrencephalic epileptic children studied by Rosvold *et al.* (1956) and Fedio and Mirsky (1969).

Three task variables were manipulated in an attempt to ascertain whether hyperactive and normal children would respond differently under different conditions. The three variables were (1) sequence, (2) distraction and (3) interstimulus interval.

It was found that the three sequences used did not differ in order of difficulty for the hyperactive and normal children, all subjects finding the *X* sequence easier than the Form or *AX* sequences at the slow (1·5 sec) interstimulus interval and the Form sequence easier than the *AX* or *X* sequences at the fast (1·0sec) interval (as measured by the absolute score).

Similarly, the second variable, namely distraction, did not differentially affect hyperactive and control children, contrary to the widely held view that hyperactive children are more susceptible to distraction than are normal children. The present study found that the performance of neither group of children was affected by the particular type of distraction used (intermittent white noise). Fedio *et al.* (1969) also failed to find any effect of an intermittent pure tone of 101 decibels on the C.P.T. performance of normal and epileptic children. Sen and Clarke (1968) point out that susceptibility to distraction is related to task difficulty. The more difficult the task, especially as it reaches the limits of the information processing capacity of the individual, the more deleterious the effect on performance of extraneous stimuli in the form of a distractor. It is possible that if the demands placed on the information processing capacities of the individual by the C.P.T. had been greater, then the intermittent noise might have produced a decline in performance. Certainly, the hypothesis that hyperactive children are more susceptible to distraction than normal children is neither confirmed nor denied by the findings of the present study but needs to be tested more systematically, taking into account such variables as type of distractor and level of task difficulty (i.e. information load).

The third variable manipulated, namely interstimulus interval, did affect the C.P.T. performance (error scores only) of hyperactive and normal children differentially, since normal children made significantly fewer errors than the hyperactive children when a slow (1·5 sec) rather than a fast (1·0 sec) speed of presentation was used. Since "errors" are responses to non-significant stimuli, this finding suggests that additional time between stimuli helped normal children more than hyperactive children to evaluate each stimulus more efficiently and thus to inhibit their responses to non-significant stimuli. That hyperactive children are more prone than normal children to impulsive responding is well documented (Campbell *et al.*, 1969; Cohen *et al.*, 1969; Conners and Greenfeld, 1966; Stevens *et al.*, 1968) and it would be interesting in a future study to see if an even longer interstimulus interval would help hyperactive children to make fewer errors on the C.P.T.

As expected, the hyperactive children were more restless (as measured by a stabilimetric cushion activity score) than the normal children. Also, their restlessness increased from the 1st to the 2nd testing session at a faster rate than did that of the normal children. While Cromwell *et al.* (1963) have suggested that the "hyperactivity" or restlessness of these children is a result of an inability to maintain attention, it is not yet possible to specify the cause-effect relationship, if any, between these two variables. As techniques for measuring activity or movement become more sophisticated, it may be possible to examine in detail the relationship between these behaviours.

The findings of the present study demonstrate that methylphenidate significantly improves the C.P.T. performance of hyperactive children: the hyperactive children on the active drug detected more significant stimuli and made fewer errors or (impulsive) responses to non-significant stimuli than did the hyperactive children given a placebo. In fact, the Drug Groups by Interstimulus Interval interactions obtained when *error* scores for the X ($p < 0{\cdot}05$) and AX ($p < 0{\cdot}10$) sequences were analysed suggests that the effect of the drug is to enable those taking it to make fewer errors at the slow speed compared to the fast speed of presentation of stimuli. In other words, it apparently helped the hyperactive children to evaluate the stimuli more effectively and thus to inhibit their responses to non-significant stimuli. This is in contrast to the hyperactive children given a placebo who did not show a similar reduction in errors at the slow interstimulus interval.

An important question which the present study could not answer owing to technical difficulties was whether methylphenidate reduced the physical restlessness of the hyperactive children. Such a reduction in the physical restlessness of hyperactive, emotionally disturbed boys given methylphenidate has recently been reported by Sprague *et al.* (1970). Clinical observation by the present investigators and the parents of the hyperactive children suggests that methylphenidate did reduce restlessness (as found by Knights *et al.*, 1969; Weiss *et al.*, 1970). In the light of the Cromwell *et al.* (1963) suggestion that the restlessness of hyperactive children is not mere activity *per se*, but continued shifting from task to task because of short attention span, it may be that the effect of methylphenidate is to improve the ability of hyperactive children to pay attention (as demonstrated on the C.P.T. in the present study) such that they shift less from task to task and thus appear less restless or active.

In conclusion, while the present study provides some evidence that hyperactive children are deficient with respect to sustained attention it is clear that future research must examine behaviours that are more closely related to the actual attentional demands placed on the child in the school setting. In other words, the emphasis must be on the discovery of the variables influencing attention in the school setting and on the manipulation of these variables in an attempt to ameliorate the attentional difficulties currently experienced in school by many hyperactive children. Also, while the present study provides some evidence that methylphenidate has a beneficial effect on the performance of certain relatively simple tasks over a short period of time, we need to know whether this beneficial effect also applies to more complex behaviours over longer periods of time.

SUMMARY

The maintenance of attention to an experimenter-paced task requiring the detection of significant stimuli was impaired in hyperactive children. When compared with a matched normal control group, the hyperactive children detected fewer of the significant stimuli and made more incorrect responses to non-significant stimuli. The presence or absence of an auditory distractor had no influence on the performance of either group of children. Those hyperactive children treated with methylphenidate (ritalin) showed a significant improvement in all aspects of their performance when compared to a control group of hyperactive children given a placebo.

REFERENCES

Baker, C. H. (1960) Observing behavior in a vigilance task. *Science* **132**, 674–675.

Cambell, S., Douglas, V. and Morgenstern, G. (1969) Cognitive styles in hyperactive children and the effect of methylphenidate. Paper read at Canadian Psychological Association Annual Meeting, Toronto.

Cohen, N., Douglas, V. and Morgenstern, G. (1969) Psychophysiological concomitants of hyperactivity in children. Paper read at Canadian Psychological Association Annual Meeting, Toronto.

Conners, C. K., Eisenberg, L. and Sharpe, L. (1964) Effects of methylphenidate (ritalin) on paired-associate learning and Porteus Maze Performance in emotionally disturbed children. *J. Consult. Psychol.* **28**, 14–22.

Conners, C. K. and Greenfeld, D. (1966) Habituation of motor startle in anxious and restless children. *J. Child Psychol. Psychiat.* **7**, 125–132.

Cromwell, R. L., Baumeister, A. and Hawkins, W. F. (1963) Research in activity level. In *Handbook of Mental Deficiency* (Edited by Ellis, N. R.), McGraw-Hill, New York.

Douglas, V., Weiss, G. and Minde, K. (1969) Learning disabilities in hyperactive children and the effect of methylphenidate. Paper read at Canadian Psychological Association Annual Meeting, Toronto.

Eisenberg, L., Conners, C. K. and Sharpe, L. (1965) A controlled study of the differential application of outpatient psychiatric treatment for children. *Japan. J. Child Psychiat.* **6**, 125–132.

Fedio, P. and Mirsky, A. F. (1969) Selective intellectual deficits in children with temporal lobe or centrencephalic epilepsy. *Neuropsychologia* **7**, 287–300.

Knights, R. M. and Hinton, G. (1969) The effects of methylphenidate (ritalin) on the motor skills and behaviour of children with learning problems. *J. Nerv. Ment. Dis.* **148**, 643–653.

Lapouse, R. and Monk, M. A. (1958) An epidemiologic study of behavior characteristics in children. *Am. J. Publ. Hlth* **48**, 1134–1144.

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, E. D. and Beck, L. H. (1956) A continuous performance test of brain damage. *J. Consult. Psychol.* **20**, 343–350.

Schulman, J. L., Kaspar, J. C. and Throne, F. M. (1965) Brain damage and behavior. *A Clinical-Experimental Study.* C. Thomas, Springfield, Illinois.

Sen, A. and Clarke, A. M. (1968) Some factors affecting distractibility in the mental retardate. *Am. J. Ment. Defic.* **73**, 50–60.

Sprague, R. L., Barnes, K. R. and Werry, J. S. (1970) Methylphenidate and thioridazine: learning, reaction time, activity, and classroom behavior in disturbed children. *Am. J. Orthopsychiat.* **40**, 615–628.

Sprague, R. L. and Toppe, L. K. (1966) Relationship between activity level and delay of reinforcement in the retarded. *J. Experiment. Child Psychol.* **3**, 390–397.

Stevens, D. A., Boydstun, J. A., Ackerman, P. T. and Dykman, R. A. (1968) Reaction time, impulsivity, and automonic lability in children with minimal brain dysfunction. *Proc. 76th Ann. Conv., Am. Psychol. Assoc.*

Stewart, M. A., Pitts, F. N., Craig, A. G. and Dieruf, W. (1966) The hyperactive child syndrome. *Am. J. Orthopsychiat.* **36**, 861–867.

Wechsler, D. (1949) *Wechsler Intelligence Scale for Children.* Psychological Corporation, New York.

Weiss, G., Minde, K., Douglas, V., Werry, J. S. and Sykes, D. H. (1970) A comparison of the effects of chlorpromazine, dextroamphetamine and methylphenidate on the behaviour and intellectual functioning of the hyperactive child. *Can. Med. Assoc. J.* In press.

Werry, J. S. and Sprague, R. L. (1970) Hyperactivity. In *Symptoms of Psychopathology* (Edited by Costello, C. G.), Wiley, New York.

Winer, B. J. (1962) *Statistical Principles in Experimental Design.* McGraw-Hill, New York.

Journal of Abnormal Child Psychology, Vol. 11, No. 4, 1983, pp. 487-496

1250

The Computerized Continuous Performance Task: A New Measure of Inattention

Steven H. Klee
Beth Israel Medical Center

Barry D. Garfinkel
University of Minnesota Medical School

Attentional performance was measured using a computerized continuous performance task, several psychometric tasks, and ratings of classroom behavior. Subjects were 51 children in the inpatient and day hospital programs of a psychiatric hospital. The relationship between performance on the computerized task and all other measures was examined. Results indicated that the continuous performance task significantly correlated with several other psychometric measures of inattention, ratings of inattention, impulsivity, and hyperactivity. The CPT had slightly better sensitivity and the same specificity as the Conners Teacher Rating Scale in identifying Conduct and Attention Deficit Disordered children. Implications for the use of the computerized continuous performance task as a screening measure for attentional difficulties is discussed.

The Continuous Performance Task (Rosvold, Mirsky, Sarason, Bransome, & Beck, 1956) was originally designed to detect and study brain damage in both children and adults. This early report indicated that the brain-damaged group's performance was consistently inferior when compared to nonimpaired controls. It was suggested that the inferior performance resulted from decreased alertness in the brain-damaged group.

Vigilance or continuous performance tasks have been used to measure sustained attention in adults (Kupietz & Richardson, 1978). Their use with Attention Deficit Disorder (ADD) children is also appropriate as these children are described as inattentive, distractible, fidgety, restless, and impulsive (Cantwell, 1972). When vigilance performance of normal

Manuscript received in final form February 23, 1983.

children was compared to that of ADD children, the index children performed less accurately (Sykes, Douglas, & Morgenstern, 1973). These authors concluded that the clinical group's poor performance was a symptom of generalized behvioral problems that appeared as inattentiveness and distractibility in other settings. Kupietz and Richardson (1978) reported that teacher ratings of classroom behavior were significantly and negatively correlated with visual vigilance task performance. Additional evidence for the relationship of poor vigilance performance and attention deficits has been provided by Douglas (1972) and more recently by Weingartner, Rapoport, Buchsbaum, Bunney, Ebert, Mikkelsen, and Caine (1980).

As continuous performance tasks are used more frequently in the investigation of attention deficits, it will become increasingly important to clarify the behavioral, perceptual, and cognitive dimensions measured by the Continuous Performance Test (CPT). Recent studies employing the CPT have replaced the laboratory equipment with more precise microcomputers. The computer is able to present stimuli at clearly defined intervals and simultaneously record subject responses. The CPT provides two dependent measures: omission and commission errors. Omission errors are reportedly sensitive to problems of inattention, while commission errors are thought to be indicative of problems of impulsivity (Sostek, Buchsbaum, & Rapoport, 1980). However, few studies have attempted to correlate CPT performance and other commonly used measures of inattention, concentration, reflectivity, impulsivity, and childhood behavior.

The present study explored the relationship between a CPT task and other psychometric measures for children hospitalized with psychiatric disorders. Vigilance as measured by CPT performance was hypothesized to correlate strongly with other psychometric measures of attention, reflectivity, and impulsivity. CPT performance was also compared to ratings of the child's behavior on the Conners Teacher Rating Scale (CTRS; Conners, 1969). It was hypothesized that poor CPT performance would identify children with marked behavioral problems as indicated by high overall scores on the CTRS. More specifically, it was predicted that CPT performance would correlated significantly with CTRS Factor 1 (conduct problems), Factor 2 (inattention), and Factor 4 (hyperactivity), and not with Factor 3 (anxiety) and Factor 5 (sociability). If the computerized CPT could identify a highly distractible and impulsive group of children, this would be a brief and accurate method of validating teacher observations and would serve to confirm deficits measured by other psychometric tests. A secondary purpose of this study was to explore the relationship of the CPT to specific diagnostic categories. Of particular interest was the usefulness of

the CPT alone and in combination with other measures in providing further clinical description of Conduct and Attention Deficit Disorder children.

METHOD

Subjects

All children admitted to the full or partial hospitalization programs at a children's psychiatric hospital constituted the subject pool. Exclusionary criteria included documented seizure disorders, brain damage, and IQ scores on the WISC-R (Wechsler, 1974) of less than 70. IQ scores were coded on a 5-point scale ranging from 1 (borderline IQ = 70-79) to 5 (superior IQ = 120+). These categories matched those provided in the WISC-R manual. The mean IQ of the subject sample was 2.7 (SD = .9) (low average-average range). The final sample consisted of 51 subjects (13 F, 38 M) with a mean age of 12.5 years (range: 7.9-16.5).

Patients involved in this study were referred for a wide variety of presenting problems, including delinquent behavior, mood disturbances, severe behavioral problems (at school and at home), and thought disorders. On the basis of DSM III diagnostic categories, subjects' admission diagnoses fell into five major categories, including Conduct Disorders (N = 23), Schizophrenic Disorders (N = 10), Attention Deficit Disorders (N = 7), Major Affective Disorders (N = 5), and Other (N = 6).

Procedure

All subjects were tested individually by one of two experimenters for 30 minutes and included the following measures:

Continuous Performance Task. This task was similar in concept to that employed by Rosvold et al. (1956). The CPT was programmed in PASCAL utilizing an Apple II computer. The task consisted of 10 letters flashing on the center of a black-and-white video monitor for 130 milliseconds at the rate of 600 milliseconds between letters. The target was the letter *S* followed by the letter *T*. The task was to press a bar when the target appeared and to avoid responding to other letter combinations. All subjects were exposed to 500 letters, with 10% of them being targets. Two dependent measures were obtained. An error of omission was scored for each target missed, while a commission error was scored when a response was made to a nontarget stimulus.

Kagan Matching Familiar Figures Test (MFFT). The children's form of the MFFT was used. This test measures reflectivity-impulsivity (Kagan, 1964), discriminates between ADD and non-ADD children, and is sensitive to stimulant drug effects (Campbell, Douglas, & Morgenstern, 1971). The test consists of 12 pictures of familiar objects followed by six variants of each picture. The subject selects from six choices the one identical to the original picture. Scoring includes latency of first response and total number of errors.

WISC-R Subtests. The Coding and Arithmetic subtests from the Wechsler Intelligence Scale for Children-Revised were administered (Wechsler, 1974). Performance on both subscales can be influenced by attention span and distractibility (Kaufman, 1979).

Behavioral Ratings. Child care workers completed the Conners Teacher Rating Scales (CTRS) on all of the children in their units. Each child care worker assigned to a particular child had the opportunity to observe the subject in both a structured classroom setting and a less structured unit setting. In order to compensate for daily fluctuations in workers' observations and subjects' behavior, rating scales were filled out at the end of the day on 3 consecutive days during the middle portion of the week. These scores were then averaged to form one score. The 3 days of rating overlapped with those of psychometric testing. The CTRS has 39 items that were scored from 0 ("not at all") to 3 ("very often"). Thus, higher total scores were indicative of more severe behavioral problems.

RESULTS

Table I shows the means, standard deviations, and ranges on the rating scale and psychometric measures for all subjects. Inspection of Table I shows that there was a wide range of performance on both the psychometric tasks and rating scales. Thus, the measures employed were sensitive in capturing the performance variability among subjects.

Table II details the correlations between CPT performance and all other measures. Correlations for errors of omission and commission were obtained separately as well as in combination (CPTOT). Consistent with Kupietz and Richardson (1978), CPT errors were found to be highly correlated with age. Therefore, partial correlations controlling for age were computed. Omission errors were positively correlated with MFFT errors and child care workers' ratings on the CTRS. High negative correlations were also found between omission errors and latency of first response on the MFFT, and scores on the WISC-R coding and arithmetic subtests. Commission errors were correlated negatively with latency and total response time

Table I. Rating Scale and Psychometric Profile

Variable	*N*	Mean	*SD*	Range
CTRS[a]	51	38.3	17.7	9.7 – 87.0
Omission errors	51	20.1	8.8	3 – 38
Commission errors	51	20.8	17.0	3 – 68
MFFT[b] (first response)	51	16.5	10.2	4.6 – 46.8
MFFT (total time)	51	22.6	11.9	7.8 – 58.5
MFFT (errors)	51	8.3	5.2	0 – 23
Coding	51	7.5	3.7	2 – 16
Arithmetic	51	8.3	2.3	4 – 13

[a]CTRS = Conners Teacher Rating Scale.
[b]MFFT = Kagan Matching Familiar Figures Test.

on the MFFT, positively with MFFT total errors and with child care workers' CTRS ratings. CPTOT was positively correlated with CTRS scores and MFFT errors, and negatively correlated with MFFT first response time and the coding subtest.

Partial correlations between CPT performance and CTRS factor scores are presented in Table III. All three measures from the computerized task were positively correlated with both Factor 2 (inattention) and Factor 4 (hyperactivity).

Though the above correlations were obtained, the relationship between CPT performance and behavioral ratings required further clarification. Subjects with a high CPT error rate were identified using a cutoff of 1/2 *SD* above the mean. As CPT and age were significantly correlated, groups were formed for younger (13.0 years and below) and older (13.1 years and above) children. Children with high error rates were chosen through comparison with their appropriate age group means. Fifteen (N = 15) children were selected as having a high error rate.

The percentage of subjects who were rated on each of the 39 behaviors as appearing "often" (2) or "very often" (3) were also noted. No item had

Table II. Correlations Between CPT Task and Other Psychometric Variables

	Partial order controlling for age						
	CTRS[c]	MFFT[d] (1st)	MFFT (total)	MFFT (errors)	Coding	Arithmetic	Age
Omission errors	.32[a]	−.28[a]	−.09	.31[a]	−.32[a]	.37[f]	−.52[b]
Commission errors	.36[b]	−.35[b]	.33[a]	.34[a]	−.25	.12	.49[b]
CPT total errors	.38[b]	−.36[b]	−.26	.28[a]	−.31[a]	.05	.56[b]

[a]$p < .05$.
[b]$p < .01$.
[c]CTRS = Conners Teacher Rating Scale.
[d]MFFT = Kagan Matching Familiar Figures Test.

Table III. Correlations Between CPT Task and Conners Teaching Rating Scale Factor Scores

	Factors				
	1 (Conduct)	2 (Inattention)	3 (Anxiety)	4 (Hyperactivity)	5 (Sociability)
Omission errors	−.05	.31[a]	−.10	.36[b]	.21
Commission errors	−.03	.33[a]	−.01	.34[a]	.05
CPT total errors	−.04	.33[a]	−.04	.35[b]	.12

[a] $p < .05$.
[b] $p < .01$.

100% endorsement. The three most endorsed CTRS items for the high CPT error group were "poor frustration tolerance" (67%), "anxious to please" (67%), and "excessive demand for attention" (60%). These items were followed by "inattentive/distractible" (47%), "short attention span" (40%), "sensitive to criticism" (40%), and "acts smart" (40%). Items occurring rarely, and therefore not characteristic of this group, were "does not get along with opposite sex" (0%), "attendance problems" (0%), "cries easily" (7%), and "shy" (7%).

One purpose of the present study was to determine the usefulness of the CPT in aiding with the clinical description and diagnosis of children. Of importance was the relative contribution of the CPT when compared to the CTRS alone. To answer this question, we selected children with scores 1/2 *SD* above the mean for both CPT performance and CTRS ratings. Sensitivity and specificity for each of these measures were determined.

Ideally, it would have been preferable to examine the sensitivity and specificity of these measures in relation to the diagnosis of ADD alone. However, as reported earlier, few subjects were hospitalized with this diagnosis. As a compromise, the ADD and Conduct Disorder groups were pooled. Thirty children met diagnostic criteria for Conduct Disorder and Attention Deficit Disorder, a group of children most likely to show deficits in concentration, impulsivity, and behavioral symptoms. Fourteen had extreme CPT scores yielding a sensitivity, or number of true positives, with this measure of 47%. By comparison, of the 21 children having diagnoses other than Conduct Disorder or Attention Deficit Disorder, 16 were not extreme responders on the CPT. Thus, specificity or number of true negatives for the CPT was 76%. The CPT was more sensitive than the CTRS, which identified 13 of the 30 Conduct Disorder and ADD children, providing a sensitivity of 43%. The specificity of the CTRS was equal to the CPT, with 76% of the other children being properly identified as not ADD or Conduct Disorder.

DISCUSSION

The usefulness of a computerized continuous performance task as a descriptive measure for attentional problems received substantial support. CPT performance was significantly correlated with other psychometric and behavioral ratings of inattention, impulsivity, and hyperactivity. A combination of omission and commission errors was found to be associated with (a) measures of impulsivity, distractibility, and lack of reflection on the MFFT; (b) overall behavioral problems as measured on the CTRS by child care workers; and (c) sustained attention as measured by the coding subtest of the WISC-R. Omission errors alone were found to correlate significantly with the arithmetic subtest, which has been viewed as a measure of sustained attention (Kaufman, 1979). Additionally, commission errors were significantly correlated with MFFT total time, which was felt to reflect the subjects' sustained attention on a task. These findings suggest that the CPT measures cognitive functions similar to other psychometric tests and behavioral ratings.

In general, the CPT was found to correlate strongly with these other measures in general, with omission scores not correlating specifically with sustained attention tasks or commission scores with measures of impulsivity alone. Rather, both of these scores correlated in a more global way with the other measures. Thus, the generally accepted notion that omission errors measure inattention and commission errors measure impulsivity was not supported in this study. Given this, the composite of CPTOT appears to be a useful summary measure in describing a subject's CPT performance.

It was expected that CPT would correlate with Factors 1 (conduct), 2 (inattention), and 4 (hyperactivity) on the CTRS. Only Factors 2 (inattention) and 4 (hyperactivity), however, correlated with CPTOT. These findings support the clinical observations that attention deficits are most often associated with symptoms of hyperactivity, impulsivity, and restlessness rather than generalized behavioral problems.

Identifying specific CTRS items characteristic of the high CPT responders provided additional support for this position. The behavioral items most endorsed by child care staff for extreme CPT responders included many of the common symptoms of Attention Deficit Disorder. Among these were poor frustration tolerance, excessive attentional demands, short attention, and high distractibility. Also, several highly endorsed items could be related to conduct disorders, including sensitivity to criticism, acting smart, and the excessive demand for attention. While CPT performance appears to be related to behaviors associated with Attention Deficit Disorders, this study was not able to demonstrate that poor

CPT performance was unique to this diagnosis. Specifically, the CPT appeared to be correlated with behaviors that might also be associated with the diagnosis of Conduct Disorder.

Limitations of the present study result in part from the sample used—namely, hospitalized children. These children are likely among the most symptomatic children in the community. As such, there is likely a great overlap of symptoms and deficits across a variety of diagnoses. Furthermore, the method of establishing diagnoses as well as the diagnostic heterogeneity of subjects were major limitations. Diagnoses were taken from admission evaluations with no attempt to validate them through independent standardized interviews. Additionally, there was no attempt to gather a minimum number of subjects in certain diagnostic categories. Rather, the current hospital census constituted the total subject pool. Therefore, there was marked overrepresentation of Conduct Disordered diagnoses, while a diagnosis of Attention Deficit Disorder was rare. Such a breakdown prevented any meaningful comparisons between these two diagnostic groups. For the present study, these diagnostic groups were pooled to provide comparisons with all other diagnostic categories.

Alone, both the CTRS and CPT had comparable moderate sensitivity and good specificity. These results indicate that with this hospitalized population, extreme responding either on the CTRS or CPT was related to a diagnosis of Conduct Disorder or ADD. A person identified as an extreme responder on either of these measures was most likely to receive a diagnosis of Conduct Disorder or ADD. Although many receiving such diagnoses were not extreme responders, a person who was either a poor CPT responder or rated poorly on the CTRS would only rarely be diagnosed as anything other than ADD or Conduct Disorder (95% specificity when combined). It should be noted that the CPT provided slightly greater sensitivity than the CTRS. It appears, therefore, to have utility as a classroom screening device with similar application as the CTRS.

The correlational data presented here are necessarily preliminary and speculative due to the nature of this study. In addition, the divisions into extreme groups using 1/2 *SD* cutoff were both retrospective and arbitrary. The value of using the CPT as a screening measure for attention deficits needs to be further addressed in future studies. Such studies should specifically address how well the CPT differentiates patients along diagnostic classifications. Comparisons among large groups of ADD, Conduct Disorder, and normal controls are needed.

In summary, the CPT was found to be a brief engaging task that can be applied to a clinical population in the preadolescent age range. The CPT appears to measure inattention and impulsivity in a global way comparable to, but not exactly the same as, the MFFT and the WISC-R Coding and

Arithmetic subtests. Similarly, it appears related to the child care workers' ratings of the child's behavior. When the results of the CPT and CTRS were combined, they yielded a good deal of diagnostic certainty. Futher study of specific behaviors that appear highly correlated to CPT performance would greatly assist teacher and clinician understanding of the child experiencing attention difficulties in the classroom. Classroom behaviors such as daydreaming, fidgeting, out-of-seat behavior, and restlessness may be better understood in the context of existing cognitive deficits as measured by this task. This study lent support to the observation that concentration and attentional problems can be easily detected on a brief screening device such as a CPT, and that these deficits are generalized and recognized by other psychometric tasks as well as behavioral ratings.

REFERENCES

Campbell, S. B., Douglas, V. I., & Morgenstern, G. Cognitive styles in hyperactive children and the effects of methylphenidate. *Journal of Child Psychology and Psychiatry*, 1971, *12*, 56-67.

Cantwell, D. P. Psychiatric illness in the families of hyperactive children. *Journal of Child Psychology and Psychiatry*, 1972, *19*, 145-153.

Conners, C. K. A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry*, 1969, *126*, 884-888.

Douglas, V. I. Stop, look and listen. The problem of sustained attention and impulse control in hyperactive and normal children. *Canadian Journal of Behavioral Science*, 1972, *4*, 259-282.

Kagan, J. *The matching familiar figures test.* Cambridge: Harvard University, 1964.

Kaufman, A. S. *Intelligent testing with the WISC-R.* New York: Wiley, 1979.

Kupietz, S. S., & Richardson, E. Children's vigilance performance and inattentiveness in the classroom. *Journal of Child Psychology and Psychiatry*, 1978, *19*, 145-153.

Rosvold, H., Mirsky, A., Sarason, I., Bransome, E., & Beck, L. A continuous performance test of brain damage. *Journal of Consulting Psychology*, 1956, *20*, 343-350.

Sostek, A. J., Buchsbaum, M. S., & Rapoport, J. L. Effects of amphetamine on vigilance performance in normal and hyperactive children. *Journal of Abnormal Child Psychology*, 1980, *8*, 491-500.

Sykes, D. H., Douglas, V. I., & Morgenstern, G. L. Sustained attention in hyperactive children. *Journal of Child Psychology and Psychiatry*, 1973, *14*, 213-220.

Wechsler, D. *Wechsler intelligence scale for children-revised.* New York: Psychological Corporation, 1974.

Weingartner, H., Rapoport, J. L., Buchsbaum, M. S., Bunney, W. E., Ebert, M. H., Mikkelsen, E. J., & Caine, E. D. Cognitive processes in normal and hyperactive children and their response to amphetamine treatment. *Journal of Abnormal Child Psychology*, 1980, *5*, 187-197.

The MCA

The MCA is a clinical research tool that has been used with considerable success to assess the effects of methylphenidate (Ritalin) in the treatment of attention deficit disorders. Although the MCA is not a proven diagnostic instrument, it has been useful when combined with classroom behavior ratings, such as the CPTQ-A and HRS, and other diagnostic procedures to help to determine whether a child might have an attention deficit disorder.

The MCA is a 25 minute psychophysiologic test using visual stimuli to measure attentional variables. This instrument has documented test-retest reliability and has been standardized for 6 to 12 year olds. The MCA is an adaptation of the VIRTEST (see Yellin, A.: A standard visual stimulus for use in studies of attention and attention deficit disorders. Res. Commun. Psychol. and Psychiat. Behav., 1980; 5:137-143 and Yellin, A., Hopwood, J. and Greenberg, L.: Adults and adolescents with attention deficit disorder. J. Clin. Psychopharm., 1982; 2:133-136).

Programmed for use with the Apple IIe, the MCA presents two easily discriminated visual stimuli in two test conditions: one in which the signal occurs infrequently in comparison to the nonsignal; in the second condition, the signal occurs more frequently than the nonsignal. The former condition, or signal infrequent, is similar to the traditional vigilance test in which errors of omission (or not responding to the signal) can be interpretd as a measure of inattentiveness. In the signal frequent condition which is a unique feature of the MCA, errors of commission (incorrectly responding to the nonsignal) can be interpreted as a measure of impulsivity or difficulty inhibiting inappropriate responses. In addition to the two types of errors, the MCA also measures response times and standard deviations which can be interpreted as an index of variability. Errors of omission, response time and standard deviations have been reported to improve when psychostimulants are administered to individuals with attention deficit disorder.

Lawrence M. Greenberg
Department of Psychiatry
University of Minnesota
Box 95
420 Delaware Street, SE
Minneapolis, MN 55455
(612) 373-8871

APPLIED INNOVATIONS, INC.

IN-SITE PSYCHOLOGICAL TESTING SOFTWARE

... *for Health Care and Human Service Professionals*

OTHER PSYCHOLOGICAL TESTING SOFTWARE AVAILABLE FROM ai

NEUROPSYCHOLOGICAL TESTING

SBORDONE-HALL MEMORY BATTERY (Revised & Updated) -- provides fully automatic serial testing of 18 discrete memory functions utilizing computer generated test stimuli.

- Produces a comprehensive quantitative and qualitative analysis of patient's performance in comparison to normative data.
- Personalized testing, understandable by all levels of users with little or no supervision required.
- Excellent for clincial assessment, cognitive rehabilitation or research.

Available on Apple II+, IIe, IIc, III (Item Number S1000) $375

THE DIGIT-DIGIT ATTENTION TEST -- fully automatic serial testing of complex attention skills utilizing randomly generated test stimuli.

- Program actually trains the individual *how* to take the test.
- Program maintains up to 900 subject files (bibliographic data, test scores, histograms of patient performance, etc.).
- Program generates various statistical and trend analyses.

Available on Apple II+, IIe, IIc, III (Item Number S1010) $250

COGNITIVE REHABILITATION

COMPLEX-ATTENTION REHABILITATION -- provides training in attentional tasks utilizing visual tracking of single or multiple stimuli.

- Interacts directly with the patient, determining the task and level of complexity.
- Advanced levels to utilize language-mediated complex attentional skills.

Available on Apple II+, II e, IIc (Item Number S1020) $195

PROBLEM-SOLVING REHABILITATION I -- provides training in visual-spatial problem-solving skills through use of a wide range of systematic tasks.
Available on Apple II+, IIe, IIc (Item Number S1030) $195.00

PROBLEM-SOLVING REHABILITATION II -- improves planning skills to facilitate effective problem-solving behavior and trains the patient to plan and evaluate the consequences of future behavior.
Available on Apple IIe, IIc (Item Number S1040) $195.00

These programs will be available soon for IBM PC and compatibles. For more information or to order, write or call us toll-free at 1-800-272-2250.

ai

Applied Innovations, Inc., South Kingstown Office Park
Wakefield, Rhode Island 02879, (401) 789-5081, (800) 272-2250

THE CONTINUOUS PERFORMANCE TASK

(Version 3.1)

***** AN OPERATING MANUAL *****

Division of Child and Adolescent Psychiatry
University of Minnesota Medical School
University Hospitals
Box 95 Mayo Memorial Building
420 Delaware Street S.E.
Minneapolis, Minnesota 55455
(612) 373-8871

TABLE OF CONTENTS

INTRODUCTION

This manual is a description of the computerized Continuous Performance Task (CPT). It is designed to assist in the administration of the CPT and provides a basic description of its program. Every attempt has been made to create a "user-friendly" system, i.e., that which requires minimal computer (specifically Apple IIe) knowledge. One will find, however, that such knowledge will aid in the understanding of the program's structure and logic. Some commonly used computer terms will be used to facilitate the brevity and clarity of this manual. It would, therefore, be advantageous to familiarize oneself wth basic computer and Apple IIe terminology.

Many users too often attempt to operate a program without first reading the accompanying manual. We recommend this not be attempted with the CPT program, since many errors will result and could lead to inappropriate parameters being set and the generation of invalid and unreliable results.

HARDWARE

The CPT software is written for specific hardware. Little deviation from these requirements will be tolerated by the software. For consistency sake, it is best to adhere to the following:

1. Apple IIe with two (2) disk drives -- non-enhanced version of ROM and microprocessor (slot #6);
2. Printer with parallel interface card (slot #1);
3. Composite color video monitor (diagonal diameter should be approximately 11");
4. 80-column card (non-extended memory; auxiliary slot #3) This card is optional but recommended (see footnote);
5. Mountain Hardware Inc. THE CLOCK™ (slot #5)
   a) The interrupt frequency required is 10 milliseconds. Necessary modifications can be performed by an individual familiar with electronics and soldering with a 25 watt soldering gun. See THE CLOCK™ operating manual (#11-00229-04) page 34;
   b) The switch on THE CLOCK™ must be in the WRITE position (see operating manual, page 3);
6. Apple joystick;
7. DOS 3.3 initialized disks (for data storage).

Footnote: Although not essential for operation of the CPT, the 80-column card allows for results to be displayed on the video-monitor in a readable fashion.

TESTING ENVIRONMENT

To obtain valid and reliable results, special care should be taken to insure a suitable testing environment. The computer should be arranged so that the center of the video monitor is at or slightly below eye level. The keyboard should be arranged below the monitor or off to the side so that the response key (see page 6) can be easily reached. The chair in which the subject sits should not be of the swivel, reclining or lounging variety; a standard desk chair would be optimal. In addition, the computer area should be clear of distractions and excessive equipment.

The room itself should be dimly lit, free of any source of light creating a glare. The room should be quiet and external noise should be minimal. If available, low level "white noise" can be present so as to block out any external and intermittent distracting activity.

As with any testing protocol, rapport should be established with the subject and any situational anxiety or oppositionality should be handled prior to the test session. Many intervening variables can affect the reliability and validity of the obtained results, as is true for many other psychometric test batteries. It is important to consider who will be administering the test. An individual who is familiar with attention deficits and how to interact with a person who presents with these deficits. A technician with no such knowledge will ineffectively instruct the subject about the test. **A person only with computer knowledge would not be an effective technician -- the subject should remain the focus and not the hardware or software.**

STARTING THE PROGRAM

If you have not already done so, check all computer connections and slot placements, specifically the clock, printer and the disk drives.

The proper disk placements are crucial. The CPT operating disk, i.e. the disk you received from the developers, must be in drive 1. A DOS 3.3 (see Apple reference manual(s) initialized disk to be used for data storage should be placed in drive 2. Obviously, make sure the drive doors are closed.

There are two ways the CPT can be accessed. One is through a warm-boot, the other is a cold-boot (see Apple reference manual). The cold-boot will result in a white band appearing at the top of the screen with the letters "CPT" superimposed. This will soon be replaced by another display with instructions for viewing an introduction text which parallels what is presented here. This section may be bypassed by pressing the [ESC] key. Scanning through the introduction is done by the ↓ and ↑ key.

Another way to access the CPT program is through the warm-boot. If this warm-boot method is used, it will result in further options being displayed on the screen after the white band described above is replaced.

Both methods, once completed, will bring the psychometrist to the **MAIN MENU** option display. This consists of four choices:

1) PRESENTATION [P]
2) SCORING [S]
3) REACTION TIMES [R]
4) EXIT PROGRAM [E]

Figure 1 displays the overall program configuration for each of these options and every section within them. This will provide the psychometrist with a reference to follow. The EXIT ([E]) option takes the psychometrist out of the CPT program; the program can be entered again at the same point by typing in RUN.

Presenting the test

If [P] is pressed, the major routine in control of presenting CPT will be loaded. This routine has two (2) subroutines, one which allows the psychometrist to enter subject demographics and test parameters, the other which accesses the main CPT program. In order to access this presentation program, the psychometrist must supply pertinent information through the demographic subroutine.

Demographic spreadsheet (option [1])

Once this portion has been loaded, the psychometrist will be presented with spaces into which information must be entered. The cursor will appear as a blinking square. Entry of information will require strict syntax. Table 1 will provide information on how this should be entered. This card indicates valid responses and examples. Only clarifications of these will follow:

NAME, ID#: If this information is not important, input of these two items can be skipped (↓ , [RETURN] or [TAB]).

SUBJ#: This must be entered in order for the data to be stored.[1]

AGE: If a child is below the age of 10, his/her age should follow a 0. (e.g. 07, "leading zero"). [see footnote -- *]

D.O.B.: Note syntax
1) Number day (e.g., 07, 31);
2) First three letters of month (e.g., NOV);
3) Last two digits of year (e.g., 84).*

SESSION: This is designed to document repeated testing. Leading zeros must be entered.[1]

GROUP: This item will set the testing inter-stimulus interval which has been derived from normative data. Ages on which this test has been normed are 7 through 12. These chronological ages are appropriate for input with leading zeros, if necessary. If a subject does not fit within this range, then either 06 or 13 are allowed. Refer to Table 2 for the resulting rates. The psychometrist may want to assign a higher "group age" to a child whose test date is within three (3) months of his/her next birthday.*

ADAPTIVE: If the adaptive version is selected, once the test begins, the starting rate will adjust according to the subject's performance. In particular, the stimulus presentation rate will increase if a correct response is made and decrease if any error is made. If the non-adaptive version is selected, the starting rate will be maintained throughout the test. We recommend the non-adaptive version.*

SERIES: Stored in memory are ten (10) one-hundred letter blocks (1-9 and A). These ten blocks have been put together to create ten different series. This feature was added to increase the flexibility of available letter sequence. Table 3 indicates these series.*

BLOCKS: The test can be given with up to ten blocks (1000 letters). Any number below can be selected. This, in conjunction with which series is selected determines which blocks will be utilized. It is recommended that the subject be given at least five (5) blocks (we prefer seven (7)) to measure in a reliable and valid manner the underlying characteristics for which the test was developed.*

START: This rate should have already been entered by the computer once the GROUP value was entered. However, this can be overwritten similar to any other supplied value within the Demographic Spreadsheet.*

MIN, MAX: These minimum and maximum rates are the limits of the interstimulus interval beyond which other perceptual variables are felt to effect the performance. These can also be changed if desired. However, if the non-adaptive version has been selected, these rates will be equal to each other. Limits of these values are .200 seconds (minimum) and 2.550 seconds (maximum).

If, after entering all necessary and desired information, the psychometrist wishes to change something, pressing [N] in response to the question:

"Information OK as shown?"

will return the cursor to the beginning of the spreadsheet. Use of the special function keys indicated on the Demographic help card will enable the psychometrist to move to the appropriate item for correction or change. Make sure that other items desired stay intact after this change (a warning message will appear if a correction is desired).

Once this information is accurate, [Y] in response to the same question the psychometrist will return to the local menu for the next selection (refer to Figure 1).

----

[1]The file name by which data will be stored and retrieved automatically by the listing:

1) The letters "CPT";
2) Four digit and/or letter subject number;
3) Session number.

EXAMPLE: CPTA20902

FOOTNOTE:

*Any invalid response will cause an error message to flash towards the bottom of the screen. Refer to the help card. Once the error has been located, the key, when pressed, will reinsert the cursor at the beginning of the item where the error occurred.

Color/letter grid (option [2])

As noted in Figure 1, this option must be selected in order for the parameters to be set. This grid is accessed by pressing [2] in response to the option menu on the screen. The psychometrist will see eight letters of different colors on the screen:

1) Dark green "U"	5) Magenta "C"
2) Purple "I"	6) Brown "L"
3) Orange "A"	7) Yellow "E"
4) White "S"	8) Dark blue "T"

If these colors do not appear as the correct color, the video monitor adjustment is essential before continuing.

At this point, the accuracy with which the subject can identify letters and colors can be determined by asking the subject to name each letter and its corresponding color. Some room should be given for personal interpretation (e.g., "red" for "magenta") but if there is strong evidence of color blindness or non-recognition of letters, the test should be stopped (as of yet, there is not a non-color version available). If the subject has no such deficits, the color version should be selected by pressing [Y]. Two target options will appear:

1) White "S"
2) Orange "S"

This refers to the color of the first letter of the target. The White 'S' (with a Blue "T") is preferred due to its clarity and display consistency on all monitors (whereas the color "orange" appears to vary in shade on different monitors). Table 4 presents the two different versions. Once a test version has been chosen, the computer will present the target letters and corresponding colors on the screen.

At this time, explanation of the test should begin and continue throughout the following sections when appropriate (refer to Testing Session Example on page 16). Once this has been done, the press of any key will return to the option menu. The psychometrist may either return to an already completed step (for further changes or clarification) — option [1] or [2] or leave the Presentation program (option [6]).

10 Letter Introduction (option [3])

This step is very useful to the subject due to its brevity and preparatory function. If selected, it will provide a ten (10) letter sequence at the rate with which the subject will see the letters during the Main Sequence. The sequence for the White "S" version is as follows:

1) Green "I"	6) Purple "H"
2) Yellow "C"	7) Orange "O"
3) Brown "E"	8) White "S"
4) White "S"	9) Green "T"
5) Blue "T"	10) Magenta "A"

If the countdown and the inter-stimulus interval is much different in rate from what was set with the Demographics, this indicates that the clock has not been properly modified for a 10 msec. interrupt. Note also that within this sequence there is a "true" target (letters #4 and #5) and a "false" target (#8 and #9) (see Explanation of Results for complete explanation (page 10)). The subject should perform perfectly during this 10 letter sequence, even if it requires multiple trials. This enables the psychometrist to be sure that the subject understands the task and is responding accurately.

NOTE: The response key is the "CLOSED APPLE" [] key. No other key on the keyboard will register any response.

100 Letter Practice (option [4])

Although optional, this sequence is highly recommended because it will indicate to the psychometrician if the subject understands the task. Clinical judgment will be necessary, since errors may either be part of a comprehension problem or a manifestation of Attention Deficit Disorder symptoms. It is esential that the psychometrist stay with the subject to answer any questions. If at any point it is necessary to stop the practice sequence due to an obvious error in comprehension, he/she may press [CTRL] [S]. Once clarification has been given, the sequence may be resumed ([R]). Because this feature is available during the Main Sequence (option [5]) and can be easily accessed by th subjects, it will be essential to hide this key sequence from the subject so he/she cannot stop the test.

The 100 letter block chosen for the practice sequence is Block #7. Note in Table 3 that this block will not appear in the main test sequence unless the test's length is over seven blocks.

Once the practice sequence is completed, the data will automatically be analyzed but not stored on disk. The indication that analysis is proceeding properly is a thick solid line building across the screen until the printing options (see Displaying Data, page 8) appear. Once the results have been scanned, basic encouragement should be given to prepare the subject for the Main Sequence.

Main Sequence -- testing session -- (option [5])

After a few preparatory steps, the countdown will appear followed by the letter sequence. During the test, the psychometrist need not stay in the room with the subject unless the subject is unduly anxious or it has been determined that his/her presence will not distract the subject. REMEMBER, THE RESPONSE KEY IS THE "CLOSED APPLE" [] KEY. NO OTHER KEY WILL REGISTER A RESPONSE.

CODE AND BLOCK DEVELOPMENT

This section explains the development of the ten 100 letter sequences. This will aid in understanding the scoring procedures.

To make this explanation easier to understand, it will be assumed that the target sequence is White "S", Blue "T". Other parameters are as follows:

1) Ten (10) targets (true target) appear in each block. This comprises twenty (20) letters;

2) Twenty (20) letter pairs appear in each block that has three of the four characteristics of the target. that is, these are "false" targets (target-like) pairs):
   a) White "S", Blue "non-T"
   b) White "S", non-blue "T"
   c) White "non-S", Blue "T"
   d) non-White "S", Blue "T"

   This comprises forty (40) letters. The fourth erroneous characteristic may be that of another target characteristic (e.g., "non-T" = "S"), but due to randomization this appears three or four times within 1000 letters;

3) The remaining forty (100-(20+40)=40) letters ("fill-in" letters) are dispersed throughout each block. These letters have none of the four characteristics of the "true" target (i.e., colors: Blue, White; letters: "S", "T");

4) No "true" target or "false" target pair will follow the same category. For example, these will never occur:
   a) White "S", Blue "T" followed by White "S", Blue "T"
   b) White "non-S", Blue "T" followed by White "non-S", Blue "T";

5) Each "true" target or "false" target pair will be followed by at least one fill-in letter, sometimes two and rarely three;

6) Linking blocks together to create the sequences will not result in any additional "true" target or "false" target pair;

7) The first letter of any block which is first in the test sequence will not be part of either a "true" target or "false" target pair. The "true" target or "false" target pair will not appear at th end of any block which is last in the test.
   See Table 10 for complete letter/color sequences for each block.

DISPLAYING DATA

Two types of data displays are available. These displays are accessible only after analysis has been completed. **It is important to note that results of analyses are not stored on disk, only raw data.** Once the psychometrist leaves the analysis routine, any clinically useful and composite results will have to be generated again through the analysis procedures (Scoring and Reaction Times). Two display modes are possible:

[1] This will display the analyzed data on the video monitor in a format which is readable (only if the 80-column board in the computer is in place). When the display has scrolled through [RETURN] will enable the psychometrist to select another display option. If the psychometrist wishes to temporarily stop the scrolling [CTRL] [S] will do so. Pressing any key will resume the scrolling once it has been temporarily stopped.

[2] (a) This option will output the analyzed results onto the printer. Once this option is selected the psychometrist will be instructed to prepare for printing and [RETURN] will activate the printer. Following printing, another [RETURN] will return the psychometrist to the local menu.

(b) An option [T] will be available to append a one page explanation (Text) of all possible errors used within the Scoring output will be readily decodable.
(The Scoring results generated and appended Text will be discussed in the section entitled Explanation of results (page 10)).

[3] This allows the psychometrist to return to the Main Menu for another selection.

ANALYSIS OF DATA

Once the presentation phase of the CPT has been completed, the data will automatically and immediately be stored under the name of CPTTEMP.DATA on the CPT system disk (drive #1). This is to insure that the immediate data would not be lost in the event of a missing or faulty data disk in drive #2 (see footnote). Once the temporary data has ben stored, it will be copied over to the data disk in drive #2 under the concatenated filename with the subject number and session number preceded by 'CPT'.

An existing or newly created data file which contains the time at which a response occurs can be scored in two different ways. Actual error classifications are not done until the analysis is completed. Specifically, analysis is completed "off-line" and not while the test sequence is being presented.

Scoring

After successfully completing this, the psychometrist may return to the Main Menu where data could be scored [S]. If a CPT test was given prior to accessing the Scoring subroutine, selection of option [0] will score this current file. If no test resides in current memory or the psychometrist wishes to score a different file, selection of option [1] will ask the psychometrist to provide a valid and existing filename.

Once this is provided, the computer will indicate which series (block sequence) was used. The following phase of analysis will take between three (3) to fifteen (15) seconds to complete, depending on the number of blocks which were used for the test session. This phase is called "number crunching" and will be indicated by the solid bright line building successive rows until all data has been analyzed.

Reaction Times

This is a supplementary analysis of the reaction times with which the subject responded and a running list fo correct responses, types of errors and when they occurred. The procedure to obtain an output uses the same procedure as that for Scoring. These results can be useful for purposes of response pattern analysis and a basic physiological measure of reaction time. Again, specific explanations will be given in Explanation of Results (page 10).

Footnote: If a disk in drive #2 is missing, the drive door is open or a duplicate filename is found, the psychometrist will see an error message and be given options to remedy the problem. If it is the case of a duplicate filename, the program will return to the subject number and session number in the Demographic spreadsheet to correct this duplication.

EXPLANATION OF RESULTS

Scoring and Text

Table 5 presents an example of the error classification printout produced by the Scoring routine. The demographic data inputted during the preparation of presenting the test is also printed and is part of the subject's permanent computer data file. Table 6 presents the Text printout which may be appended to the error classification printout only if option [2] of DISPLAYING DATA (page 8) is selected. The following section will explain the types of responses recorded. The section CODE AND BLOCK DEVELOPMENT should be used for reference to accompany this explanation. In addition, a few terms need to be defined:

A) 'True' target: This is a two letter sequence presented in appropriate colors (i.e. White 'S', Blue 'T') which if responded to would preclude an Omission.

B) 'False' target: A two letter sequence which has only three (3) characteristics of the 'true" target (see CODE AND BLOCK DEVELOPMENT, page 7).

C) window: CPT scoring is based on a letter interval paradigm. Each occurrence of a letter marks the start of another time interval which is determined by the setting of rate (start, min, max) parameters. The rate is also an inter-stimulus interval. A 'true' or a 'false' target response is scored in relation to the number of intervals that have been presented since the onset of the second letter of the 'true' or 'false' target. The operational definition of an open (or within) window response is a response occurring within two intervals following the onset of the second letter of a 'true' or 'false' target. A closed (or outside) window response is a response occurring after two intervals following the onset of the second letter of a 'true' or 'false' target, but before the first letter of another 'true' or 'false' target has been presented.

The error classification is presentd in three sections. The first section is a composite profile of the subject's overall performance:

RESPONSES: This is the number of key presses recorded while the subject was responding correctly or incorrectly. If for each block, the subject responded only to a correct target, the number "10" would appear.

CORRECT: This represents the number of 'true' targets to which the subject responded correctly. Since there are exactly ten (10) targets per block, a percentage value can be derived.

ERROR: This section presents the number of errors, regardless of the type which the subject committed. This section is further broken down in the second section of the printout.

The second section classifies errors according to the major categories -- Omissions and Commissions.

OMISSION: These numbers and percentages will obviously be the arithmetic complement of the number of correct responses. An operational definition of an Omission is as follows:

> **A failure to press the 'closed apple' [] key during the open window rsponse interval (within two intervals).**

LATE OMISSION: This second type of Omission error is scored if the subject responds but during a closed window. The number of these errors possible is dependent on which blocks are being presented and how many "fill-in" letters occur between 'true' and 'false' targets. Table 7 presents the maximum number of the late omission responses possible for each block. Refer to COMMISSION TYPE III-A for a differential classification of these errors.

COMMISSION TYPE I: The error is registered when an open window response occurs in reaction to any of the four 'false' targets (Refer to CODE AND BLOCK DEVELOPMENT). The third section of this printout expands these responses to the frequency with which each 'false' target response occurred.

COMMISSION TYPE II: This Commission error category is similar in definition to the LATE OMISSION (applied to closed window responses) but indicates the frequency of a 'false' target rsponse. Table 7 presents the maximum number of these commission errors possible. Differential classificiations are possible of these errors. (Refer to COMMISSION TYPE III-A).

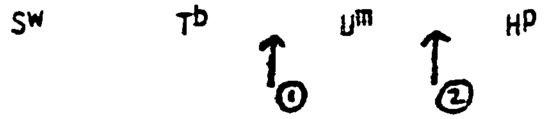
COMMISSION TYPE III: This category can provide important information as to whether the test can be considered valid. In formulating this error classification the goal was to provide a categorization of errors specific to each behavioral response. The following categories can provide crucial information regarding the subject's comprehension of the test, anticipatory tendencies and over-responsivity to the stimulus separate from the COMMISSION TYPE I and II errors:

COMMISSION TYPE III-A: This error category includes all responses made to the first letter of either a 'true' or 'false' target before the second letter appears. These are classified an Anticipatory Response errors. These responses may ocur in individuals who require more time than is given within the response window. Therefore, these responses may be a reflection of the number of "fill-in" letters before the anticipatory responses and/or the neurological intactness of the subject.

These errors may more appropriately be LATE OMISSION or COMMISSION TYPE II. If these errors occur frequently, the Reaction Time scoring routine can provide additional information regarding the feasibility of an anticipatory response in relation to a minimum physiologica speed with which to decode a stimuli. Refer to Reaction Time from further explanation (page 13).

COMMISSION TYPE III-B: These errors will be recorded if a response is made in a subsequent letter interval although the subject responded in the preceding interval. These are called Additional Response errors. These errors can be generated by responding either to a 'true' or 'false' target. These types of errors can be subdivided:

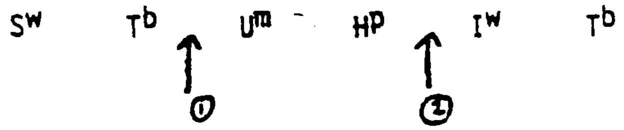
Open window: An additional response (as described above) during the allowed two letter interval which is operationally used as the active response window. For example, with the arrows indicating a response and the superscript letter indicating the color of the letter:



The first rsponse would be recorded as a correct response to a target. The second response would be recorded as an open window.

COMMISSION TYPE III-B, since the latter response occurred before the onset of the purple 'H' which marks the close of the target window. If the 'S' was of a different color, the second response would stay the same, but the first response would be considered as a COMMISSION TYPE III-D (See CODE AND BLOCK DEVELOPMENT).

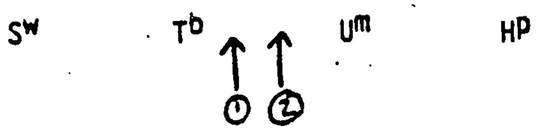
Closed window: This definition logically follows the previous error. An example should suffice:



The second response is after the active response window has closed but before the onset of a letter which is part of a 'false' target (White non-S, Blue 'T').

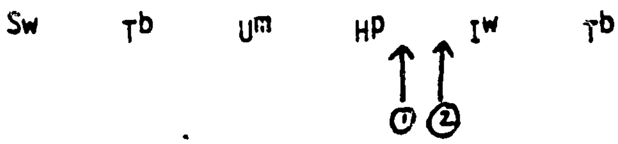
COMMISSION TYPE III-C: This error will be recorded when a subject responds two or more times within the same letter interval. This is referred to as a Multiple Response error:

Open window: This category will be chosen for a response such as this:



would generate a correct response and a COMMISSION TYPE III-C open window error.

Closed window: This response style:



would be described as one LATE OMISSION and one COMMISSION TYPE III-C closed window response. If the target was instead, a 'false' target, the first response would be considered a COMMISSION TYPE II error; the error category for the second response would remain the same.

If either of the two previous categories (COMMISSION TYPE II and III) occur with a relatively high frequency, one may speculate a variety of task inappropriate behaviors ranging from poor comprehension to oppositionality to extreme impulsivity. Although these errors alone should not invalidate the test, consideration should be given to the degree of off-task behavior alters the maningfulness of the test and the behavioral problems it is designed to measure.

Reaction Times

At first glance, the Reaction Time printout looks incomprehensible. However, if one remembers that the CPT scoring is based on a letter interval response paradigm, appropriate definitions begin to fall in place.

Any display of the reaction time data will include the demographic data. The columns numbered 0 through 16 will appear before the first response of each block administered. These numbers correspond to the response categories in the Scoring prinout (see Table 5 and 9). For each response made, an asterisk will appear below one category (0-16) preceded by five values which can be used to pinpoint when the response was made. STIM denotes the nth letter in which the response occurred.

All other categories (EVENT, cTIME, RATE, tTIME) indicates values read from THE CLOCK". Each of the numbers signify the number of 10 milliseconds time intervals for that category. For simplicity, since the test begins at time = 0, multiplying the number by 10 indicates the time that has elapsed. More specific explanations for these categories will make the scoring routine more clear:

RATE: This indicates the length of the letter interval in which the response occurred (40 = 400 milliseconds). This value, if using the non-adaptive version of the test, should be the same as the starting, minimum and maximum rates set in the demographic spreadsheet and should remain the same through the entire test.

cTIME: This value indicates the tie by which the interval ends for the letter in which a response occurred. As shown in Table 8, this person responded for the first time during the sixth letter whose time interval closes at cTIME = 240 (2400 milliseconds from the beginning of the test).

EVENT: This is simple the time when the actual response occurred. Therefore, knowing cTIME and STIM, EVENT identifies exactly where within the STIM interval the response was made.

tTIME: This value is the length of time that has elapsed since the onset of the first letter of either a 'true' and 'false' target given that both of the letters of either target has been presented. More specifically, if tTIME is positive, this indicates that a response was made after the second leetter of either a 'true' or 'false' target. If, however, tTIME is negative, this mans that a response was made before the onset of the second letter (i.e., Anticipatory). The value indicates the length of time still to elapse before the onset of the second letter. Any positive value over 120 (1200 milliseconds) indicates a Late Response. Obviously, to obtain a reaction time, one needs to subtract 40 from tTIME to determine the elapsed time since onset of the second letter of either 'true' or 'false' target. The formula for Reaction Time is therefore:

RT = 9tTIME * 10) - (RATE * 10)

Illustrations of these examples will **clarify** the **Reaction Time** printout. See **Table 8 (Response A).**

A ---	stimulus#	**4**	**5**	**6**	**7**	**8**
	$letter^{color}$	O^m	S^w	T^b	U^o	Y^w
	Interval close time(cTIME)		**160**	**200**	**240**	
	Response (STIM) = 6)				↑	
	Response time (EVENT)				224	
	tTIME = 64					

In response to the Blue 'T' of a **'true' target,** the subject responded 240 milliseconds into the Blue 'T' interval (**RT** = 640 - 400). An asterisk appears in column 0 of **STIM** - row **6**.

B ---	stimulus#	**311**	**312**	313	314	315	316
	$letter^{color}$	P^p	S^w	T^b	E^o	H^g	U^w
	Interval close time(cTIME)		12440	12480	12520	12560	12600
	Response (STIM)						↑
	Response time (EVENT)						12570
	tTIME = 131 (RT = 1310 - 400)						

C ---	stimulus#	265	266	267	268	269
	$letter^{color}$	S^w	T^b	L^y	S^g	T^b
	Interval close time(cTIME)		10600	10640	10680	10720
	Response (STIM)					↑
	Response Time (EVENT)					10704
	tTIME = -16 (Anticipatory response Type 9)					

In this case, the usual RT formula cannot be used (whenever **tTIME** is negative). Instead the length of time that must elapse until the onset of the second ltter is **l tTIME l** * 10 (absolute value of **tTIME** * 10 = 160 milliseconds).

NOTE: This is a case where a differential classification may be speculated. The sequence before the response included a 'true' target followed by one "fill-in" letter. It is possible that the subject required more time in which to respond correctly than the sequence allowed. This re-classification should be made on the basis of the responses the subject gave throughout the entire test. If **anticipatory** errors occur frequently following a target with only one "fill-in" letter, this indicates that the subject needs more time in which to respond (>RATE * 2 milliseconds). If such a response follows a 'false' target, this may suggest immediate or 'delayed' impulsivity.

TESTING SESSION EXAMPLE

The following describes a sample testing session. This may be used as a model by which to guide oneself through the test instructions given to the subject and appropriate responses to screen displays. U will signify the psychometrist dialogue; S will signify the subject dialogue. D will signify the screen display and I will signify information typed in on th keyboard.

Prior to the subject entering the room, the Presentation program has been loaded [P] and subsequently the Demographic Spreadsheet [1]. Demographic information has been entered and verified as correct. The parameters set for data storage and test administration are:

	Display	Input
1)	Subject#:	A101
2)	Sex:	M
3)	Session:	01
4)	Group:	10
5)	Adaptive:	N
6)	Series:	01
7)	Blocks:	07
8)	Start:	already set at .600 due to input as a 10 year old male
9)	Min,Max:	already set at .600 due to input as the non-adaptive rate version.

D: "Information OK as shown?"
I(U): [Y]
D: ** CPT PRESENTATION **
[1] Enter DEMOGRAPHICS
[2] Color/letter GRID
[3] EXIT to main menu

Perform which function? [X]

I(U): [2]
D: Eight colored letter appear in a 3X3 grid
"Color version (Y or N)?" [X]

The subject is brought in. If the psychometrist does not know the subject, a few minutes should be taken to establish rapport, which should include obtaining information about his/her familiarity with computers. The subject should be comfortably sitting in a chair in front of the computer video monitor. Equipment rearrangement may be necessary.

U: "This is a game to see how well you can pay attention to what will appear on the screen. But first, I'd like to find out if you can see letters and colors okay.. this is a green 'U', right?" (point to upper left letter). "What's this?" (point to upper middle letter; repeat this procedure as necessary*).
I(U): [Y] (If subject identifies the letters and colors adequately)
D: Same eight letter grid with:
[W] White-S
[O] Orange-S
Which option: ☒

I(U): [W]
D: A White 'S', Blue 'T' appears in the middle of the screen
"Press ANY KEY to RETURN to menu" ☒

U:** "I'd like to explain exactly what you have to do for this game. You will have a chance to practice, so don't worry if you're not sure about what it is you're supposed to do. Feel free to ask any questions about the game itself. It's very important that you ask questions if you have any, because it wouldn't be fair to you if I'm asking you to do something that you're confused about. Many children have questions about this game, so please don't feel that you're dumb if you don't understand something I said, okay?"
S: "What's this thing?" (points to printer, if in sight)
U: "It's a machine that types things out, but you won't be using it. The only thing that you have to use is one key on this keyboard while you're looking at this screen. Once we get started, you'll be seeing letters flashing on the screen one right after another. And the letters will be in different colors, too. So, one letter will flash on the screen (a hand motion to demonstrate a 'flash' may be helpful), then that letter will go off but another will come on right after that."

---

* The psychometrist may correct the subject after he/she has completed the identification phase. A determination will have to be made whether or not these errors warrant discontinuation of the test.
** This, of course, is an exampl of a smooth introduction. The psychometrist should feel free to deter from this presented method according to the needs of the subject.

"Sometimes you may see a White 'S' (point to the screen) and then right after that a Blue 'T' may appear. If you see these two letters appearing one right after the other, in these colors and in this order (frequently refer to the target appearing on the screen), prss this key (point to the 'closed apple' [] key). It's important that you press only this key because the computer won't know that you'r paying attention if you press other keys."

S: "What if I see an 'S' and a 'T' in different colors?"

U:* "That's a very good question! The only time you want to press this key is if the 'S' is a White 'S' and the 'T' is a Blue 'T'. The computer will try to trick you by showing you all different combinations of letters and colors which are very close to the White 'S', Blue 'T', but pressing the key for anything other than the White 'S', Blue 'T' would be a mistake. If you do make a mistake, don't get mad or frustrated with yourself -- everyone makes some mistakes. Just do the best you can."

S: "Do I press this button after the White 'S'?

U:* "No, it's important to press this button (point) after the Blue 'T' only if the White 'S' comes before it. Just because you see a White 'S' doesn't mean you should press this button; you should wait to see if the Blue 'T' comes next because the computer may try to trick you like I mentioned before. Why don't you practice a bit to show me that you understand what you're supposed to do."

I(U): (Any key pressed will advance the program)

D: (The CPT Presentation menu has appeared again with three additional options -- [3], [4], and [5])

I(U): [3] (10 letter INTRODUCTION)

D: "Loading Display Routine" (3 seconds)
(New Screen) "Press ANY KEY to BEGIN Series"

I(S): Subject presses a key and sequence starts.

U: "Always let me set the game up; you'll have a chance to play in a second" (The sequence ends).

D: "Press ANY KEY to RETURN to menu"

I(U): (Any key is pressed followed by option [3])

U: "I'm going to show you 10 letters one right after another. What are you supposed to do when you see the different colored letters?"

S: "I should press this key (points to the 'closed apple [] key) whenever I see a White 'S' and a Blue 'T'."

---

* Make sure to include these explanations even if the subject does not ask the above question. This is crucial to the subject understanding the test.

U: "That's right. Press this key if you see a Blue 'T' only when a White 'S' came before it. Good! In a few seconds, you'll see a countdown with numbers which tells you to get ready. Then some letters will start flashing. There may be a White 'S', Blue 'T' in there; if so, press right after you see the Blue 'T'," (Make sure the subject is prepared; have subject put finger right above the 'closed apple' key as if ready to press)
I(U): (Any key is pressed)
D: (Countdown starts from the number 5 and continues at the rate at which the psychometrist requested.
I(U): Subject presses both for the 'true' and 'false' target. The 'false' target response is followed by "Oops".
U: "The computer caught you! See, that's how the computer will try to trick you. Try your best to pay close attention.* Okay, let's try it again (Return to 10 letter introduction option and begin sequence).
I(U): The subject responds to the 'true' target correctly and appropriately inhibits a response to the 'false' target.
U: Excellent! You did a perfect job! Now, let's have some more practice. (Return to menu and choose option [4] -- 100 Letter Practice)
D: "Loading code for practice sequence" followed by
"Press ANY KEY to BEGIN Practice"
U: "You'll see a few more letters, now. Ready?
I(U): (A key press begins the sequence)
U: (Observe the subject's responses to determine comprehension and attitude toward the test. Discuss blatant errors such as multiple or random responses whether or not they're associated with the stimuli and return to previous steps where appropriate) "Okay, let's see how you did."
D: "Analyzing Practice Results" (The practice sequence will automatically be analyzed; consecutive solid lines will be forming).

[1] Display Results on 80 col. MONITOR
[2] Output Results to PRINTER
[3] EXIT to local menu

---

* If it is evident that the subject is waiting a long time to respond correctly or if he/she hesitates to press because one letter has gone by after the Blue 'T' (i.e. the window is still open), encourage him/her to respond as soon as he/she can.

I(U): [2] (It is best to have a permanent copy of the practice session -- practice results are not stored on disk). Printout shows nothing indicative of non-compliance or poor comprehension. Therefore, continue
U: "Okay, now we're ready for the real game."
I(U): Return to option menu [RETURN] then [3] for EXIT to Local menu and select [5] -- Main Sequence.
D: "Loading Series #1"
"Press ANY KEY to BEGIN sequence"
U: I'm going to leave the room (or "I'll be back here") so I won't disturb you. Do the best you can and remember, White S, Blue T (sequence begins).
D: (Subject leaves following completion of the test sequence)
"**PLEASE DO NOT DISTURB**"
"STORING CPTTEMP.DATA on sys disk"
"STORING CCPTA10101 on data disk"*
(New Screen)
"FILE:CPTA10101 ALREADY EXISTS"
"for SUBJECT #A101"
SESSION #01
"Program WILL NOT overwrite"
"Enter RETURN to correct demographics"
I(U): [RETURN]
D: "Loading demographic spreadsheet" (program will return cursor to SUBJECT # and SESSION to allow for correction).
I(U): (Changes SUBJECT # to A102 and [TAB] through SESSION)
D: "Information OK as shown?"
I(U): [Y]
Another possible error which may occur at this point is:
D: "Storing CPTA10101 on data disk"
(New Screen)
"I/O ERROR"
"CHECK:"
"1) Is the CPT-SYSTEM Disk in D1?"
"2) Is the DATA Disk in D2 INITIALIZED?"
"3) Are the Drive Doors SHUT?"
"Please correct the problem"
"Then press RETURN to continue"
I(U): (Drive #2 door was not shut) [RETURN]
D: "Storing CPTA10101 on Data Disk

---

* The following steps will not occur unless there has been an error in how the test results will be stored. Refer to the step marked by the check mark ( ).

( ) (New Screen)
"***CPT PRESENTATION***"
[1] Enter DEMOGRAPHICS
[2] EXIT to main menu
"Perform which function?" [X]
I(U): [2] (If the psychometrist wishes to administer another test,
[1] will lead back to the point where the TESTING SESSION EXAMPLE began).
D: (New screen) -- "CPT Main Menu"
I(U): [S] -- Scoring
D: "Loading Scoring Routine"
"***CPT SCORING***"
[0] SCORE Current File CPTAA10201
[1] LOAD New file
[2] EXIT to main menu
"Perform which function? [X]
I(U): [0]
D: "Loading series #1"
"Busy number crunching (solid line extends across screen)"
(New Screen)
"[1] Display Results on 80 col. MONITOR
[2] Output Results to PRINTER
[3] EXIT to local menu"
"Perform which function?" [X]
I(U): [2], [RETURN] (Printer prints out results similar to Table 5)
D: "[T] Append explanation TEXT
[R] RETURN to menu"
"Perform which function? [X]
I(U): [T] (Printer will print out Table 6)

After TEXT is printed, the computer will return to the printing options. [3] will return to the Scoring options; a [2] selection will return to the Main Menu, where the psychometrist can set up for another Presentation, Scoring, Reaction Time or EXIT. If desired, typing in RUN will return the psychometrist to the point where the program was exited.

REFERENCES AND BIBLIOGRAPHY

Garfinkel BD, Klee SH (1983) A computerized assessment battery for attention deficits. The Psychiatric Hospital 14(3):163-166.

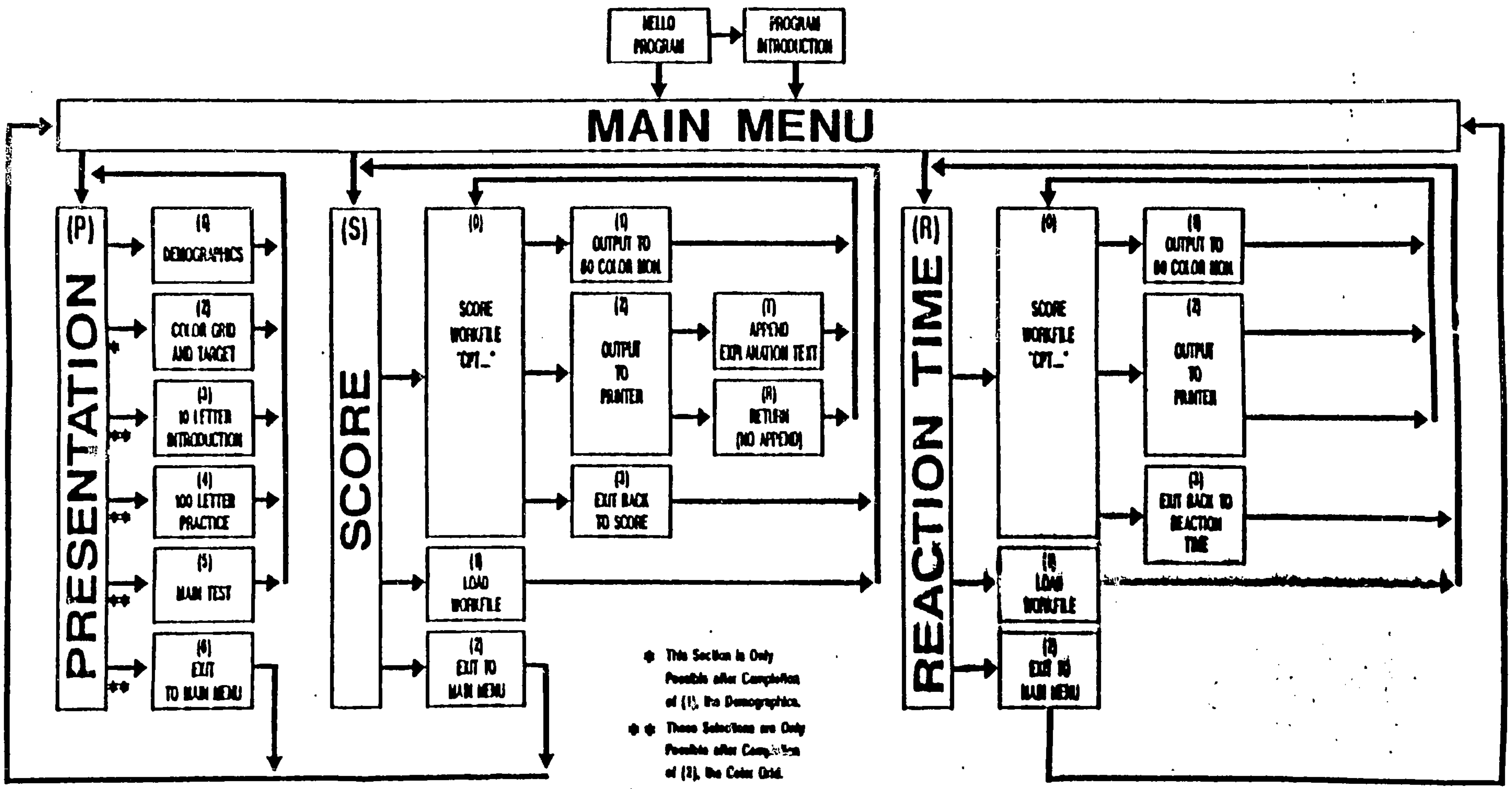
Garfinkel BD (1984) Neuroendocrine and cognitive responses to amphetamine in adolescents with a history of ADD. In: Attention Deficit Disorder: Diagnostic, Cognitive and Therapeutic Understanding. (Ed.) L. Bloomingdale, Spectrum Publications Inc.: New York.

Klee SH, Garfinkel BD (1983) The computerized continuous performance task: A new measure of inattention. Journal of Abnormal Child Psychology 11(4):487-496.

Shapiro SK, Garfinkel BD (submitted, 1985) The occurrences of behavior disorders in children: The interdependence of Attention Deficit Disorder and Conduct Disorder.

#164, SS-SSB

FIGURE 1



HELLO PROGRAM
PROGRAM INTRODUCTION
MAIN MENU
(P) PRESENTATION
(1) DEMOGRAPHICS
(2) COLOR GRID AND TARGET
(3) 10 LETTER INTRODUCTION
(4) 100 LETTER PRACTICE
(5) MAIN TEST
(6) EXIT TO MAIN MENU
(S) SCORE
(0) SCORE WORKFILE "CPT_"
(1) OUTPUT TO 80 COLOR MON.
(2) OUTPUT TO PRINTER
(1) APPEND EXPLANATION TEXT
(2) RETURN (NO APPEND)
(3) EXIT BACK TO SCORE
(1) LOAD WORKFILE
(2) EXIT TO MAIN MENU
(R) REACTION TIME
(0) SCORE WORKFILE "CPT_"
(1) OUTPUT TO 80 COLOR MON.
(2) OUTPUT TO PRINTER
(3) EXIT BACK TO REACTION TIME
(1) LOAD WORKFILE
(2) EXIT TO MAIN MENU
* This Section is Only Possible after Completion of (1), the Demographics.
** These Selections are Only Possible after Completion of (2), the Color Grid.

TABLE 1
DEMOGRAPHIC HELP CARD

Field	Valid Responses	Number of Digits	Examples; Comments	Notes
Name	Letters, numbers, punctuation	24	Smith, John Henry	
ID#	Letters, numbers, punctuation	7-1	D134729-2; Hospital or Clinic ID	
Subj#	Letters, numbers	4	A209; Subject number	1
Age	00,01,02........ 99	2	30; Biological Age (years)	
Sex	M or F	1	M	
D.O.B.	01 to 31--3 letters--00 to 99	2-3-2	30-AUG-54; Date of Birth	2
Date	01 to 31--3 letters--00 to 99	2-3-2	20-SEP-84; Test Date	2
Session	00,01,02.........99	2	02	1
Group	06,07,08........13	2	13; Group reflects the test age	3,4
Adaptive	Y or N	1	Y=Adaptive Rate; N=Non-adaptive Rate	5
Series	01,02,03.........10	2	Each series is a fixed, random combination of ten blocks.	
Blocks	01,02,03..........10	2	Each block contains 100 letters. This entry determines the length of the test.	
Start	0.200 to 2.550	4	Rate (in seconds) at which letters appear	4,5
Min	0.200 to 2.550	4	Fastest rate attainable	4,5
Max	0.200 to 2.550	4	Slowest rate attainable	4,5

Special Function keys:

- Right Arrow: Moves the cursor right one character position.
- Left Arrow: Moves the cursor left one character position.
- Down Arrow: Moves the cursor to the first character position of the next field.
- Up Arrow: Moves the cursor to the first character of the previous field.
- Return: Same as Down Arrow.
- Tab: Same as Down Arrow.

Notes:

1) The subject and session fields are combined to form the file name. Using the examples from each field respectively, the created file name would be: CPTA20902.
2) **Do not use** 8-30-54 for 30-AUG-54. The computer will not accept it.
3) The **group age** (test age) corresponds to chronological age in years; round to higher age when test date is within three months of next birthday.
4) Age appropriate rates are automatically inserted according to the value of the **Group** (Refer to **Table 2**: Default Presentation Rates).
5) If the non-adaptive selection is chosen (i.e., N) then **Min** and **Max** will automatically be set to the same rate as **Start**.

TABLE 2

DEFAULT PRESENTATION RATES*

Group (years)	Male (seconds)	Female (seconds)
< 7	.990	.720
7	.990	.720
8	.700	.640
9	.670	.640
10	.600	.640
11	.530	570
12	530	.520
13	.400	400
> 13	.400	400

* These will be the rates at which the letters are presented at the beginning (with the adaptive rate) or throughout the test. They can be changed if desired within the **Demographic Spreadsheet.**

TABLE 3

SEQUENCE OF BLOCKS FOR EACH SERIES

Series	Sequence of Blocks
1	4 9 1 6 2 5 A 7 3 8
2	5 3 9 8 2 4 1 A 7 6
3	8 2 6 4 1 3 A 4 7 5
4	9 5 3 A 1 4 8 6 2 7
5	2 A 5 1 6 8 3 7 4 9
6	4 8 2 5 3 A 1 6 7 9
7	8 A 4 6 1 3 9 2 5 7
8	5 1 6 3 9 4 8 A 7 2
9	9 3 6 4 1 5 A 7 2 8
10	4 A 8 1 6 9 5 2 3 7

TABLE 4

COLOR USAGE FOR CPT VERSION

	Test Version*	
Color Name	White S Blue T	Orange S Blue T
Magenta	X	X
Dark Blue	X	X
Purple	X	X
Dark Green	X	X
Brown	X	
Orange	X	X
Pink		X
Green		X
Yellow	X	X
White	X	

* Colors used for the two target options available.

These steps were made to change the White S, Blue T version to the Orange S, Blue T:

1) Orange --------> White
2) Pink ----------> Orange
3) Green ---------> Brown

TABLE 5
Scoring printout

Name: Doe, John   ID#: 11111111   Subject #: A101
Age: 25 Sex: M DOB: 21-FEB-59 Test Date: 04-FEB-85 Session #: 01
Group: 13 Adaptive: N Series#: 01 # of Blocks: 04 Color: W
Starting Rate: 0.400 Minimum Rate: 0.400 Maximum Rate: 0.400

Block	Responses #	Correct #	Correct %	Error #
1	13	8	80%	7
2	7	6	60%	5
3	12	9	90%	4
4	12	9	90%	3
Total	44	32	80%	19

Block	Errors						
				'Commission Type' Errors			
	Omissions #	Omissions %	Late Omissions #	TYPE I #	TYPE I %	TYPE II #	TYPE III #
1	2	20%	0	4	20%	0	1
2	4	40%	0	1	5%	0	0
3	1	10%	0	2	10%	0	1
4	0	0%	1	2	10%	0	0
Total	7	18%	1	9	11%	0	2

Block	Type I (expanded)								Type II (expanded)			
	A #	A %	B #	B %	C #	C %	D #	D %	A #	B #	C #	D #
1	3	60%	0	0%	0	0%	1	20%	0	0	0	0
2	0	0%	1	10%	0	0%	0	0%	0	0	0	0
3	2	40%	0	0%	0	0%	0	0%	0	0	0	0
4	0	0%	2	20%	0	0%	0	0%	0	0	0	0
Total	5	25%	3	15%	0	0%	1	5%	0	0	0	0

Block	Type III (expanded)						
	A			B		C	
	1	2	3	1	2	1	2
1	0	0	0	1	0	0	0
2	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0
4	0	0	0	0	0	0	0
Total	0	0	1	1	0	0	0

TABLE 6
TEXT printout

TERMS USED

A) (TRUE) TARGET: White 'S' immediately followed by Blue 'T'

B) FALSE TARGET: Two letter sequence having only three (3) characteristics of true target (e.g., Green 'S', Blue 'T'; see TYPE I: A-D below)

C) WINDOW: Two-letter interval beginning from the onset of the second letter of a true or false target

1) OPEN WINDOW: less than two-letter intervals have passed

2) CLOSED WINDOW: Two letter intervals have passed

DEFINITIONS FOR ITEMS SCORED

--- CORRECT: Open window response to target

--- OMISSION errors: No response to target

--- LATE OMISSION errors: Closed window response to target (*)

--- COMMISSION TYPE I errors: Open window response to a false target (*)

A) White 'S', Blue non-'T'
B) White 'S', non-Blue 'T'
C) White non-'S', Blue 'T'
D) non-White 'S', Blue 'T'

--- COMMISSION TYPE II errors: Closed window response to a false target

A - D) Same as for 'TYPE I errors'

--- COMMISSION TYPE III errors: Additional errors

A) Response to first letter of a true or false target

1) White 'S'
2) White non-'S'
3) non-White 'S'

B) An additional response within a different letter-interval than an already recorded response

1) Open window
2) Closed window

C) An additional response within the same letter-interval as an already recorded response

1) Open window
2) Closed window

* The definition of these errors appear incorrectly on the computer generated printout

TABLE 7

MAXIMUM NUMBER OF CLOSED WINDOW RESPONSE ERRORS POSSIBLE*

Block	Late Omissions	Commission Type II
1	2	5
2	1	6
3	5	3
4	3	5
5	3	5
6	2	7
7	5	3
8	5	3
9	4	4
A	3	4

* Determined by random placement of "fill-in" letters occuring after a 'true' or 'false' target.

TABLE 8

Reaction Time printout

Name: Doe, John ID #: 11111111 Subject #: A101
Age: 25 Sex: M DOB: 21-FEB-59 Test Date: 04-FEB-85 Session #: 01
Group: 13 Adaptive: N Series #: 01 # of Blocks: 04 Color: W
Starting Rate: 0.400 Minimum Rate: 0.400 Maximum Rate: 0.400

STIM	EVENT	cTIME	RATE	tTIME	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
6	224	240	40	64	•																
12	467	480	40	67		•															
13	498	520	40	98						•											
21	811	840	40	51	•																
31	1230	1240	40	70	•																
38	1505	1520	40	65					•												
45	1785	1800	40	65	•																
51	2031	2040	40	71	•																
58	2303	2320	40	63	•																
67	2647	2680	40	87		•															
79	3152	3160	40	72	•																
86	3435	3440	40	75	•																
89	3552	3560	40	72		•															

STIM	EVENT	cTIME	RATE	tTIME	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
104	4123	4160	40	83			•														
146	5826	5840	40	66	•																
161	6434	6440	40	74	•																
172	6870	6880	40	70	•																
185	7391	7400	40	71	•																
191	7632	7640	40	72	•																
198	7911	7920	40	71	•																

STIM	EVENT	cTIME	RATE	tTIME	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
204	8150	8160	40	70	•																
213	8505	8520	40	65	•																
230	9190	9200	40	70	•																
235	9375	9400	40	95		•															
243	9711	9720	40	71		•															
247	9874	9880	40	74	•																
259	10353	10360	40	73	•																
268	10704	10720	40	-16										•							
272	10878	10880	40	78	•																
279	11135	11160	40	95	•																
285	11361	11400	40	81	•																
293	11716	11720	40	76	•																

STIM	EVENT	cTIME	RATE	tTIME	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
303	12118	12120	40	78	•																
315	12571	12600	40	131											•						
323	12912	12920	40	72			•														
336	13439	13440	40	79	•																
342	13677	13680	40	77	•																
348	13912	13920	40	72	•																
355	14191	14200	40	71	•																
362	14468	14480	40	68	•																
368	14716	14720	40	76	•																
375	14988	15000	40	68	•																
389	15555	15560	40	75	•																
400	15961	16000	40	81			•														

TABLE 9

ERROR CODES USED IN REACTION TIME PRINTOUT

Code*	Type of Error	Description
0	(CORRECT)	Punctual response to White S/Blue T
1	I-A	to White S/Blue non-T
2	I-B	to White S/non-Blue T
3	I-C	to White non-S/Blue T
4	I-D	to non-White S/Blue T
5	III-B-1	Additional response (within window)
6	III-C-1	Multiple response (within window)
7	III-A-1	Response to White S
8	III-A-2	Response to White non-S
9	III-A-3	Response to non-White S
10	LATE	Late response to White S/Blue T
11	II-A	to White S/Blue non-T
12	II-B	to White S/non-Blue T
13	II-C	to White non-S/Blue T
14	II-D	to non-White S/Blue T
15	III-B-2	Additional response (outside window)
16	III-C-2	Multiple response (outside window)

* These codes correspond to the codes which will appear in the **Reaction Time** printout.

TABLE 10 (A)

Block 1

	Orange U
	Yellow I
TGT	White S, Blue T
	Yellow L
	Green Y
B	White S, White T
	Purple F
	Brown A
	Magenta I
TGT	White S, Blue T
	Purple L
C	White P, Blue T
	Yellow U
	Brown U
B	White S, Yellow T
	Purple U
	Magenta I
D	Magenta S, Blue T
	Magenta L
B	White S, Orange T
	Orange C
TGT	White S, Blue T
	Yellow I
	Green F
A	White S, Blue P
	Orange H
D	Green S, Blue T
	Purple P
C	White U, Blue T
	Brown C
A	White S, Blue L
	Orange H
	Orange E
TGT	White S, Blue T
	Purple U
C	White C, Blue T
	Green O
A	White S, Blue L
	Magenta U
B	White S, Brown T
	Purple F
TGT	White S, Blue T
	Green P
C	White P, Blue T
	Magenta U
	Green F
TGT	White S, Blue T
	Yellow L
D	Green S, Blue T
	Magenta I
TGT	White S, Blue T
	Brown U
B	White S, White T
	Magenta I
TGT	White S, Blue T
	Magenta U
A	White S, Blue Y
	Brown Y
TGT	White S, Blue T
	Magenta L
D	Blue S, Blue T
	Magenta E
C	White A, Blue T
	Green I
TGT	White S, Blue T
	Yellow A
A	White S, Blue U
	Orange U
D	Purple S, Blue T
	Purple H

Block 2

	Yellow U
	Magenta I
C	White Y, Blue T
	Magenta E
TGT	White S, Blue T
	Purple O
B	White S, Orange T
	Brown F
	Green L
C	White E, Blue T
	Orange A
TGT	White S, Blue T
	Green A
A	White S, Blue P
	Brown C
	Green H
TGT	White S, Blue T
	Magenta Y
D	Brown S, Blue T
	Purple F
TGT	White S, Blue T
	Brown Y
D	Yellow S, Blue T
	Orange P
	Magenta L
C	White I, Blue T
	Green Y
TGT	White S, Blue T
	Magenta A
D	Blue S, Blue T
	Yellow U
C	White C, Blue T
	Orange O
B	White S, Purple T
	Brown E
C	White A, Blue T
	Yellow Y
A	White S, Blue H
	Orange U
	Green O
B	White S, Purple T
	Orange H
TGT	White S, Blue T
	Purple U
A	White S, Blue I
	Magenta H
B	White S, White T
	Magenta I
	Brown C
TGT	White S, Blue T
	Purple F
A	White S, Blue P
	Brown H
TGT	White S, Blue T
	Brown P
D	Brown S, Blue T
	Purple P
B	White S, Yellow T
	Green P
	Purple U
D	Yellow S, Blue T
	Magenta H
TGT	White S, Blue T
	Green Y
A	White S, Blue O
	Purple C
TGT	White S, Blue T
	Magenta F
	Yellow U
	Yellow E

TABLE 10 (B)

Block 3

TGT	White S, Blue T
	Yellow U
A	White S, Blue H
	Orange C
TGT	White S, Blue T
	Green E
	Magenta I
D	Purple S, Blue T
	Yellow C
TGT	White S, Blue T
	Brown U
	Yell-- I
	Green H
C	White C, Blue T
	Brown I
TGT	White S, Blue T
	Yellow H
B	White S, Magenta T
	Orange H
D	Magenta S, Blue T
	Magenta F
B	White S, Orange T
	Purple I
TGT	White S, Blue T
	Orange E
	Green L
D	Brown S, Blue T
	Purple P
A	White S, Blue P
	Pur--- H
D	Magenta S, Blue T
	Brown Y
A	White S, Blue P
	Orange U
TGT	White S, Blue T
	Purple C
C	White L, Blue T
	Yellow I
	Magenta Y
	Orange E
A	White S, Blue E
	Orange F
C	White H, Blue T
	Brown P
	Yellow I
TGT	White S, Blue T
	Green H
B	White S, Purple T
	Orange A
A	White S, Blue U
	Brown U
B	White S, White T
	Orange I
TGT	White S, Blue T
	Green E
C	White H, Blue T
	Magenta U
TGT	White S, Blue T
	Purple L
	Green O
B	White S, Purple T
	Orange C
	Orange I
TGT	White S, Blue T
	Yellow H
	Orange Y
C	White O, Blue T
	Green A
D	Orange S, Blue T
	Purple A

Block 4

	Brown B
C	White H, Blue T
	Magenta O
TGT	White S, Blue T
	Orange U
C	White Y, Blue T
	Yellow H
A	White S, Blue Y
	Purple H
TGT	White S, Blue T
	Orange B
B	White S, Brown T
	Yellow F
TGT	White S, Blue T
	Orange P
C	White Y, Blue T
	Magenta P
	Yellow P
D	Brown S, Blue T
	Green L
TGT	White S, Blue T
	Green I
	Orange F
B	White S, Brown T
	Magenta C
D	Orange S, Blue T
	Brown C
C	White Y, Blue T
	Green U
	Yellow Y
TGT	White S, Blue T
	Orange L
A	White S, Blue S
	Purple L
TGT	White S, Blue T
	Yellow C
C	White P, Blue T
	Brown I
	Green C
TGT	White S, Blue T
	Brown O
	Magenta E
	Purple H
B	White S, Magenta T
	Yellow I
A	White S, Blue S
	Purple P
B	White S, Orange T
	Brown P
	Yellow L
D	Brown S, Blue T
	Magenta H
B	White S, Purple T
	Magenta E
TGT	White S, Blue T
	Orange P
	Brown P
D	Green S, Blue T
	Green Y
TGT	White S, Blue T
	Brown C
A	White S, Blue C
	Yellow E
TGT	White S, Blue T
	Purple I
D	Orange S, Blue T
	Magenta H
A	White S, Blue F
	Yellow L
	Brown P

TABLE 10 (C)

Block 5

	Magenta O
D	Orange S, Blue T
	Green F
	Yellow H
TGT	White S; Blue T
	Magenta P
	Orange Y
	Green H
B	White S, Magenta T
	Magenta O
TGT	White S, Blue T
	Brown F
D	Purple S, Blue T
	Brown Y
A	White S, Blue L
	Brown A
TGT	White S, Blue T
	Green C
C	White H, Blue T
	Purple P
TGT	White S, Blue T
	Brown Y
C	White Y, Blue T
	Orange E
B	White S, Magenta T
	Magenta O
D	Yellow S, Blue T
	Green E
TGT	White S, Blue T
	Purple O
	Brown A
A	White S, Blue E
	Yellow H
TGT	White S, Blue T
	Green L
D	Orange S, Blue T
	Brown P
B	White S, Yellow T
	Purple C
A	White S, Blue A
	Yellow H
D	Orange S, Blue T
	Purple C
A	White S, Blue F
	Green F
C	White Y, Blue T
	Purple U
TGT	White S, Blue T
	Brown E
A	White S, Blue I
	Green Y
	Purple P
TGT	White S, Blue T
	Magenta C
B	White S, Purple T
	Orange A
	Brown I
TGT	White S, Blue T
	Green A
C	White E, Blue T
	Magenta U
	Yellow E
B	White S, Brown T
	Orange O
C	White A, Blue T
	Yellow I
	Purple E
TGT	White S, Blue T
	Green Y
	Orange I

Block 6

	Purple I
TGT	White S, Blue T
	Brown A
A	White S, Blue I
	Yellow I
	Purple H
D	Brown S, Blue T
	Purple P
TGT	White S, Blue T
	Orange E
	Green H
C	White U, Blue T
	Magenta Y
A	White S, Blue O
	Magenta L
B	White S, Purple T
	Yellow A
A	White S, Blue A
	Orange F
D	Purple S, Blue T
	Brown U
	Green F
C	White Y, Blue T
	Purple U
TGT	White S, Blue T
	Yellow H
C	White I, Blue T
	Green P
TGT	White S, Blue T
	Brown O
D	Blue S, Blue T
	Yellow P
TGT	White S, Blue T
	Orange A
B	White S, Green T
	Yellow C
	Purple Y
TGT	White S, Blue T
	Brown I
A	White S, Blue E
	Magenta Y
	Green L
TGT	White S, Blue T
	Yellow O
B	White S, Yellow T
	Orange U
TGT	White S, Blue T
	Brown C
	Magenta F
C	White F, Blue T
	Yellow U
TGT	White S, Blue T
	Yellow H
A	White S, Blue Y
	Green F
	Brown I
B	White S, Purple T
	Yellow A
	Brown H
D	Orange S, Blue T
	Magenta E
TGT	White S, Blue T
	Brown O
C	White T, Blue T
	Purple E
D	Green S, Blue T
	Purple O
	Purple P
B	White S, White T
	Orange I

TABLE 10 (D)

Block 7

- C White H, Blue T
  Yellow L
- B White S, Green T
  Magenta P
- D Brown S, Blue T
  Brown U
- TGT White S, Blue T
  Yellow U
  Yellow H
- B White S, Purple T
  Orange P
- A White S, Blue I
  Green E
- B White S, Orange T
  Purple U
- TGT White S, Blue T
  Yellow U
  Green I
- B White S, White T
  Orange L
- TGT White S, Blue T
  Brown U
- C White H, Blue T
  Magenta L
- A White S, Blue E
  Green F
- D Green S, Blue T
  Purple P
  Orange H
- A White S, Blue A
  Yellow I
- TGT White S, Blue T
  Brown A
- C Green S, Blue T
  Yellow O
- D White O, Blue T
  Magenta U
  Green Y
- TGT White S, Blue T
  Orange I
  Purple U
- C White P, Blue T
  Yellow E
  Magenta I
- A White S, Blue H
  Green C
- TGT White S, Blue T
  Purple A
- A White S, Blue H
  Magenta L
- TGT White S, Blue T
  Yellow F
  Green I
- C White I, Blue T
  Brown A
- TGT White S, Blue T
  Green U
  Purple F
  Yellow H
- D Magenta S, Blue T
  Brown I
- B White S, Yellow T
  Magenta A
- TGT White S, Blue T
  Orange E
- D Yellow S, Blue T
  Purple L
- TGT White S, Blue T
  Green U
  Yellow [illegible]

Block 8

- Brown C
- D Yellow S, Blue T
  Green U
- C White T, Blue T
  Orange E
- TGT White S, Blue T
  Magenta H
  Purple Y
- B White S, Magenta T
  Brown U
- TGT White S, Blue T
  Brown Y
- A White S, Blue C
  Green P
- TGT White S, Blue T
  Green F
- A White S, Blue O
  Yellow F
- TGT White S, Blue T
  Green U
- D Purple S, Blue T
  Yellow A
  Yellow O
- C White Y, Blue T
  Magenta H
  Purple E
- B White S, Magenta T
  Magenta F
- C White L, Blue T
  Purple F
- B White S, Orange T
  Orange A
- TGT White S, Blue T
  Purple Y
  Brown I
- C White A, Blue T
  Purple U
- A White S, Blue C
  Magenta L
- TGT White S, Blue T
  Magenta E
  Yellow A
- C White C, Blue T
  Purple P
- TGT White S, Blue T
  Orange O
- D Brown S, Blue T
  Green C
- A White S, Blue S
  Orange H
- TGT White S, Blue T
  Yellow A
- D Blue S, Blue T
  Magenta O
- B White S, Green T
  Green O
- D Orange S, Blue T
  Yellow H
- TGT White S, Blue T
  Green P
  Magenta U
  Purple C
- B White S, Purple T
  Yellow O
  Orange F
- A White S, Blue U
  Yellow L
- TGT White S, Blue T
  Yellow H
  [illegible]

TABLE 10 (E)

Block 9

Green O
B White S, Orange T
Purple I
Yellow C
Magenta A
TGT White S, Blue T
Orange L
Yellow O
A White S, Blue O
Brown C
TGT White S, Blue T
Yellow H
D Magenta S, Blue T
Green P
A White S, Blue C
Yellow U
Brown O
D Purple S, Blue T
Green Y
B White S, Green T
Green E
A White S, Blue E
Purple Y
TGT White S, Blue T
Magenta U
B White S, Green T
Purple H
TGT White S, Blue T
Green C
C White F, Blue T
Magenta C
TGT White S, Blue T
Purple U
C White C, Blue T
Yellow U
B White S, Magenta T
Magenta L
C White L, Blue T
Orange P
A White S, Blue S
Magenta E
TGT White S, Blue T
Green E
Yellow C
C White I, Blue T
Orange P
Magenta A
D Orange S, Blue T
Orange O
TGT White S, Blue T
Purple I
A White S, Blue P
Orange E
C White L, Blue T
Purple F
Green P
D Green S, Blue T
Yellow O
TGT White S, Blue T
Purple O
B White S, Magenta T
Yellow A
TGT White S, Blue T
Brown H
Magenta Y
D Yellow S, Blue T
Orange F
TGT White S, Blue T
Purple O
Green I

Block A (10)

D Orange S, Blue T
Yellow H
C White E, Blue T
Green A
Magenta L
Brown U
TGT White S, Blue T
Yellow C
Orange O
Green C
B White S, White T
Yellow L
TGT White S, Blue T
Purple H
C White F, Blue T
Orange P
TGT White S, Blue T
Yellow U
Brown U
C White I, Blue T
Green O
D Yellow S, Blue T
Orange O
Brown F
C White S, Blue C
Orange F
TGT White S, Blue T
Orange H
D Blue S, Blue T
Orange U
B White S, Magenta T
Orange F
TGT White S, Blue T
Purple E
B White S, Orange T
Green Y
Magenta L
A White S, Blue O
Brown E
Green E
Yellow F
TGT White S, Blue T
Purple U
C White C, Blue T
Purple E
A White S, Blue I
Orange O
B White S, Orange T
Magenta O
D Yellow S, Blue T
Brown H
A White S, Blue L
Orange U
TGT White S, Blue T
Orange I
A White S, Blue E
Magenta F
B White S, Green T
Orange F
TGT White S, Blue T
Orange P
C White Y, Blue T
Green C
TGT White S, Blue T
Yellow L
D Yellow S, Blue T
Yellow Y
TGT White S, Blue T
Orange E
Brown A

STANDARDIZATION OF THE GORDON DIAGNOSTIC SYSTEM

The Gordon Diagnostic System (GDS) was standardized on 1356 children. This current sample is comprised of 716 boys and 640 girls 3 to 16 years of age. There are 306 3-5 year olds, 337 6-7 year olds, 263 8-9 year olds, 221 10-11 year olds and 229 12-16 year olds.

The children were randomly selected from class lists from both public and private schools. According to the Hollingshead measure of socioeconomic status, subjects came from the full range of social classes. The children were primarily from the greater Syracuse area as well as from Virginia and Philadelphia. The children were administered the GDS on an individual basis in school by trained examiners. A child's scores were included in the standardization sample if the child had no known psychiatric problems, learning problems, had not been retained, had never received psychoactive medication, and showed no evidence of frank neurological damage.

GDS scores have been found unaffected by SES or geographic location. Comparison with the norms of an inner-city sample of impoverished children and a sample from an impoverished rural area revealed no significant differences in mean scores.

There currently is a standardization study being conducted by Dr. Jose Bauermeister with Puerto Rican children. The sample contains approximately 200 children who were randomly selected from class lists of public and private schools. Another large project is being conducted with Department of Defense dependents in Okinawa by Sandra Sacco.

Analyses of standardization data have revealed no significant sex differences. Consequently, identical norms are presented for boys and girls.

As indicated in the GDS Interpretive Supplement (page 1), the Threshold Tables divide the standardization sample into three groups according to statistical convention in the behavioral sciences. As such, a child's score is considered ABNORMAL if it falls in the 5th percentile or lower. The BORDERLINE range includes children with scores in the 6th - 25th percentile, and the NORMAL range is for those with scores above the 25th percentile.

INFORMATION
REGARDING
STANDARDIZATION

WIDE RANGE ACHIEVEMENT TEST-REVISED

LEVEL 1

AGE NORMS

AGE: 11-6 to 11-11

Stand. Score	RAW SCORE Reading	Spelling	Arithmetic	Percentile	Stand. Score	RAW SCORE Reading	Spelling	Arithmetic	Percentile
46	43	—	—	.03	101	77	49	33	53
47	44	24	—	.04	102	78	—	—	55
48	—	—	16	.05	103	—	—	—	58
49	45	25	—	.06	104	79	50	34	61
50	—	—	—	.07	105	80	—	—	63
51	46	—	17	.08	106	—	51	—	66
52	47	26	—	.09	107	81	—	35	68
53	—	—	—	.1	108	—	52	—	70
54	48	27	18	.2	109	82	—	—	73
55	49	—	—	.3	110	83	53	36	75
56	—	28	—	.4	111	—	—	—	77
57	50	—	19	.5	112	84	54	—	79
58	—	29	—	.6	113	85	—	37	81
59	51	—	—	.7	114	—	55	—	82
60	52	30	20	.8	115	86	—	—	84
61	—	—	—	.9	116	—	—	38	86
62	53	31	—	1	117	87	56	—	87
63	54	—	21	1	118	88	—	—	88
64	—	—	—	1	119	—	57	39	90
65	55	32	—	1	120	89	—	—	91
66	—	—	22	1	121	—	58	—	92
67	56	33	—	1	122	90	—	—	93
68	57	—	—	2	123	91	59	40	94
69	—	34	23	2	124	—	—	—	95
70	58	—	—	2	125	92	60	—	95
71	—	35	—	3	126	93	—	41	96
72	59	—	—	3	127	—	61	—	96
73	60	36	24	4	128	94	—	—	97
74	—	—	—	4	129	—	—	42	97
75	61	37	—	5	130	95	62	—	98
76	62	—	25	5	131	96	—	—	98
77	—	—	—	6	132	—	63	43	98
78	63	38	—	7	133	97	—	—	99
79	—	—	26	8	134	98	64	—	99
80	64	39	—	9	135	—	—	44	99
81	65	—	—	10	136	99	65	—	99
82	—	40	27	12	137	—	—	—	99
83	66	—	—	13	138	100	—	45	99
84	67	41	—	14	139	—	—	—	99.1
85	—	—	28	16	140	—	—	—	99.2
86	68	42	—	18	141	—	—	46	99.3
87	—	—	—	19	142	—	—	—	99.4
88	69	43	29	21	143	—	—	—	99.5
89	70	—	—	23	144	—	—	47	99.6
90	—	—	—	25	145	—	—	—	99.7
91	71	44	30	27	146	—	—	—	99.8
92	—	—	—	30	147	—	—	—	99.9
93	72	45	—	32	148	—	—	48	99.91
94	73	—	31	34	149	—	—	—	99.92
95	—	46	—	37	150	—	—	—	99.93
96	74	—	—	39	151	—	—	49	99.94
97	75	47	—	42	152	—	—	—	99.95
98	—	—	32	45	153	—	—	—	99.96
99	76	48	—	47	154	—	—	50	99.97
100	—	—	—	50	155	—	—	—	99.98

THE GORDON DIAGNOSTIC SYSTEM

GDS RECORD FORM

Name: .... Date Tested: ....

Sex: .... Date of Birth: ....

Age: ....

Referred by: ....

Presenting Problem:

Duration: ....

School: .... Grade: ....

Special Academic Programming:

Remarks:

CLINICAL DIAGNOSTICS INC
1536 Cole Blvd., Golden, Colorado 80401
Suite 350 • 303-232-7129

GDS RECORD FORM

DELAY TASK

TASK PARAMETERS

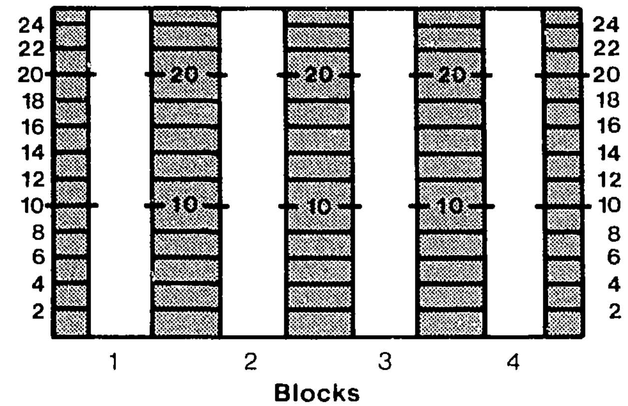
☐ Standard — Delay Interval = 6 Seconds; Block Length = 120 Seconds

☐ Other Fixed Parameters — Delay Interval .... Seconds; Block Length .... Seconds

☐ Variable Parameters

Block	1	2	3	4
Delay Interval	....	....	....	....
Block Length	....	....	....	....

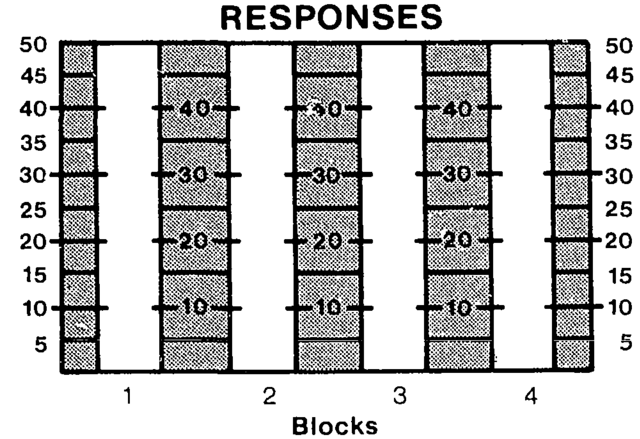
REWARDS



RESULTS

Data Selector Positions	Display		
1	....	Total:	Rewards
2	....		Responses
3	....	Block 1:	Rewards
4	....		Responses
5	....	Block 2:	Rewards
6	....		Responses
7	....	Block 3:	Rewards
8	....		Responses
9	....	Block 4:	Rewards
10	....		Responses

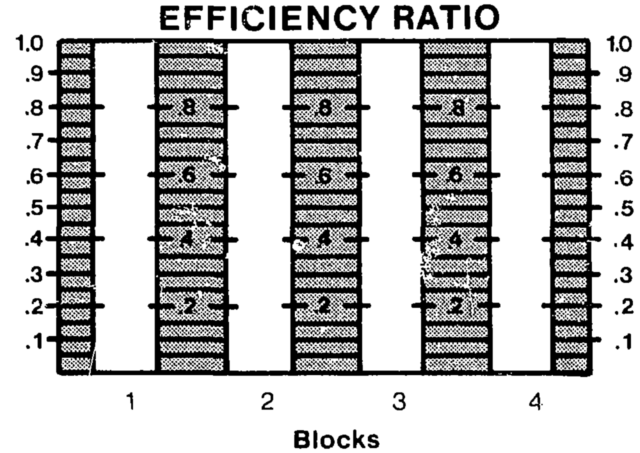
RESPONSES



SUMMARY STATISTICS

Block	1	2	3	4	Total
Interval					
Length					
Rewards					
Responses					
Effic. Ratio					

EFFICIENCY RATIO



STRATEGIES/COMMENTS:

GDS RECORD FORM

VIGILANCE TASK

TASK PARAMETERS

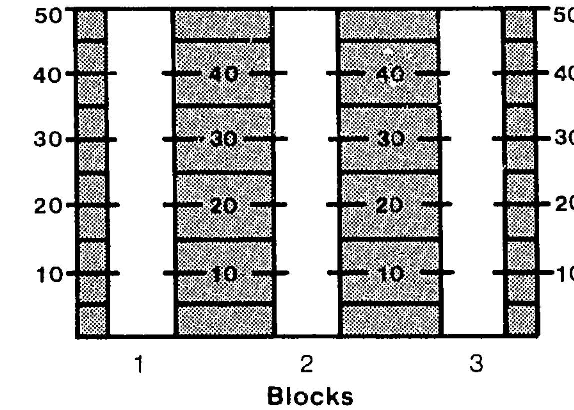
☐ Standard — Pres. Interval = 1 Second; Block Length = 180 Seconds

☐ Other Fixed Parameters — Pres. Interval .... Seconds; Block Length .... Seconds

☐ Variable Parameters

Block	1	2	3
Presentation Interval	....	....	....
Block Length	....	....	....

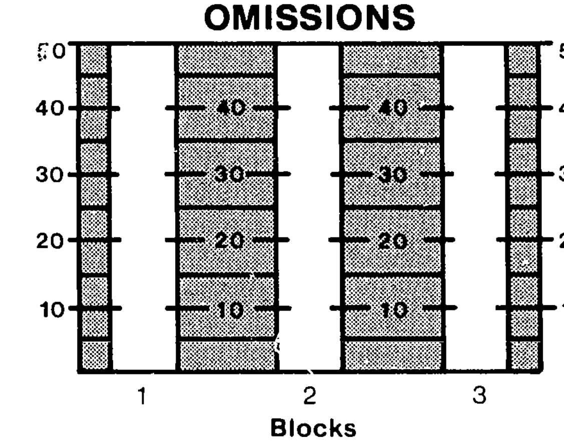
CORRECT



RESULTS

Data Selector Positions	Display		
1	....	Total:	Correct
2	....	Block 1:	Correct
3	....		Omissions
4	....		Commissions
5	....	Block 2:	Correct
6	....		Omissions
7	....		Commissions
8	....	Block 3:	Correct
9	....		Omissions
10	....		Commissions

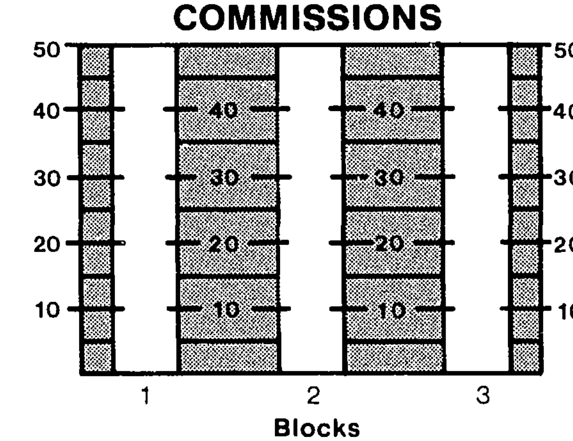
OMISSIONS



SUMMARY STATISTICS

Block	1	2	3	Total
Interval				
Length				
Correct				
Omissions				
Commissions				

COMMISSIONS



STRATEGIES/COMMENTS:

GDS RECORD FORM

ADDITIONAL DATA

WISC-R Full Scale IQ ....
Verbal IQ ....
Performance IQ ....
Freedom from Distractibility Factor ....

RATING SCALES Connors Hyperactivity Factor ....
CBCL Hyperactivity Factor ....
Other: ....

BENDER Koppitz Score .... Age Equivalent ....
Emotional Indicators ....
Significant Test Behaviors
(Neurogenic Compensations) ....

OTHER SOURCES OF INFORMATION

Classroom Observations:

Teacher Reports/Physician Reports:

Medical, Social or Developmental Reports:

DIAGNOSTIC IMPRESSIONS:

COMPARING CLASSROOM AND CLINIC MEASURES OF ADD-H: DOSE RESPONSE EFFECTS

MARK D. RAPPORT, GEORGE J. DUPAUL, KEVIN KELLY, AND JAY JONES

THE CHILDREN'S LEARNING CLINIC
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF RHODE ISLAND
KINSTON, RI 02881

PAPER PRESENTED AT THE AMERICAN PSYCHOLOGICAL ASSOCIATION
LOS ANGELES, CALIFORNIA, AUGUST, 1985

Drug Studies

Abstract

Attention Deficit Disorder with Hyperactivity (ADD-H) is characterized by an inability to sustain attention, impulsivity, and overactivity (DSM-III; American Psychiatric Association, 1980). Related problems such as social and academic difficulties are frequently reported by parents and teachers (Aman, 1984). Although some follow-up studies have shown a reduction in untreated ADD-H children's symptomatology over the course of three to four years (Charles & Schain, 1981; Satterfield, Satterfield, & Cantwell, 1981), a corresponding improvement in academic functioning and social adjustment has not been demonstrated (Satterfield, Hoppe, & Schell, 1982).

The attentional component of the disorder is especially problematic owing to its interferring effects on school performance, home behavior and persistence into adult life, and has been empirically validated in both laboratory (Sykes, Douglas, & Morgenstern, 1973) and classroom settings (Rapport, Murphy, & Bailey, 1980; 1982). Thus far, the Continuous Performance Test (CPT; Rosvold, Mirosky, Sarason, Bransome, & Beck, 1956) has been the most frequently used laboratory test for identifying the attentional deficiencies in ADD-H children (Klee & Garfinkel, 1983). The CPT is an experimenter-paced vigilance task sensitive to momentary lapses of attention and yields two dependent measures: omission and commission errors; indicative of attentional and impulsivity problems, respectively (Sostek, Buchsbaum, & Rapoport, 1980). ADD-H children typically make fewer correct detections (i.e., omission errors) and more incorrect responses (i.e., commission errors) to presented stimuli compared to normal children.

Psychostimulant medication is presently the most widely used treatment for ADD-H (Safer & Krager, 1983), most likely due to its cost-efficiency and

demonstrated short-term effects on sustained attention (Sykes et al., 1973), activity level (Porrino et al., 1983), academic performance (Rapport, Stoner, DuPaul, Birmingham, & Tucker, 1985), and impulsivity (Brown & Sleator, 1979; Rapport, DuPaul, Stoner, Birmingham, & Masse, in press). Despite the findings of increased attention in ADD-H children in both school (Rapport et al., 1982; 1985) and laboratory (Sostek et al., 1980) environments following psychostimulant treatment, there is a dearth of information as to whether drug-related changes in these two settings are similar and/or dose specific.

The primary purpose of the present study was to investigate the relationship between the most frequently used laboratory method of assessing attentional deficits (CPT) and ADD-H children's attention in school under various doses of methylphenidate. A second purpose was to subject the expected dose-response curves to statistical analysis and determine whether a similar quadratic relationship (inverted U-shaped curve) is found between performance and dose as reported in previous studies (Brown & Sleator, 1979; Sprague & Sleator, 1977). Finally, a comparison of fixed-dose and milligram- per-kilogram (mg/kg) dose-response curves was planned to determine the relative advantage of each in assessing drug responsivity. Collectively, these results should have a direct bearing on assessing ADD-H children's responsiveness to psychostimulant medication and the putative ability of laboratory-based testing for establishing optimal dose.

METHOD

Subjects

Families of children with complaints of inattentiveness, impulsivity, and overactivity were referred to the Children's Learning Clinic (CLC) of the University of Rhode Island by pediatricians and local school system personnel. All children met each of the following diagnostic criteria: (a) an independent diagnosis by the child's pediatrician and CLC's directing clinical psychologist using DSM-III (American Psychiatric Association, 1980) criteria for Attention Deficit Disorder with Hyperactivity (ADD-H); (b) a maternal report of a developmental history consistent with ADD-H (Barkley, 1981); (c) a maternal rating of at least two standard deviations above the mean for the child's age on the Werry-Weiss-Peters Activity Scale (WWPAS: Routh, Schroeder, & O'Tuama, 1974) and problems in at least 50% of the situations on Barkley's (1981) Home Situations Questionaire; (d) a teacher rating on the Abbreviated Conners Teacher Rating Scale (ACTRS) above 15, the designated cutoff score for hyperactivity (Werry, Sprague, & Cohen, 1975); (e) performance on the Matching Familiar Figures Test (MFFT: Kagan, Rosman, Day, Albert, & Phillips, 1964) characteristic of a "fast-inaccurate or impulsive" responder (i.e., faster than average responses and higher than average error rates for the child's age); and (f) absense of any gross neurological, sensory, or motor impairment as determined by pediatric examination.

A total of 24 children (20 males, 4 females) between the ages of 7 and 10 years (M = 8.8 years) participated in the study after informed consent was obtained from their parents. All children were of average or above average intelligence (mean IQ = 98.3; SD = 8.8) as assessed by the Peabody Picture Vocabulary Test-Revised, Form L (Dunn & Dunn, 1981) and from families of low to middle socioeconomic status (Hollingshead, 1975). No children were taking medication prior to participating in the study.

Experimental Design and Drug Administration

A double-blind, placebo control, within-subject (crossover) experimental design was used in which all children received each of the five doses (following baseline measures) in a randomly assigned, counterbalanced sequence.

Methylphenidate (MPH) was prescribed by each child's pediatrician in the following doses: placebo, 5 mg, 10 mg, 15 mg, and 20 mg. Fixed doses (vs mg/kg) were prescribed to reflect typical pediatric practice in the United States (Physicians' Desk Reference, 1985) and because children's response to MPH dosage manipulations has not been shown to be dependent upon total body weight (Kinsbourne & Swanson, 1979, p. 14; Rapport et al., 1985). All MPH and placebo dosages were packaged in colored gelatin capsules by the clinic's pharmacist to avoid detection of dose and taste. Capsules were placed in individual daily dated envelopes to insure accurate dose administration.

Children were seen once per week at the CLC for individual testing. Baseline measures were obtained during the child's second clinic visit, to allow familiarization with clinic personnel and testing procedures. On subsequent testing days, children were administered a capsule of either the active agent (MPH) or a placebo. Testing (under double-blind conditions) was begun 90 min after oral ingestion to insure optimal medication effect (cf. Kinsbourne & Swanson, 1979, p. 11).

Children were classified as favorable responders or non-responders to MPH in a manner similar to previous studies (Thurston, Sobol, Swanson, & Kinsbourne, 1979), using the Paired Associates Learning (PAL) test. The PAL test was administered once per week under each experimental condition described above. Those children showing a drug-induced facilitation in performance of 25% of more (i.e., a 25% drop in error rate) compared to baseline or placebo were classified as favorable responders, while those whose performance neither improved to the above criteria nor declined in relation to baseline or placebo were classified as non-responders. All children in the present study were found to be favorable responders using the above criteria.

Assessment

Clinic Measures

During each child's weekly session the Gordon Diagnostic System (GDS; Gordon & McClure, 1983) version of the Continuous Performance Test (CPT) was administered. In using this version, the child is presented with a series of numbers ranging from one through nine at 1 sec intervals and is required to press a button on the apparatus each time the number "1" is immediately followed by the number "9". Total correct responses, omission (failing to identify the "1-9" combination when it appears) and commission errors (pressing the button for digit combinations other than the correct combination) are recorded for the 9 min test duration.

Classroom Measures

The classroom phase of the experiment involved 6 consecutive weeks of assessment and observation (see Teacher Ratings and Behavioral Observations sections below) which corresponded with the 6 week clinic testing period decribed above. Following baseline data collection (1st week), parents were given a week's worth of medication in pre-dated envelopes at one dose level (i.e., placebo, 5 mg, 10 mg, 15 mg, or 20 mg MPH). This prodecure continued until a child received each dose for 6 consecutive days. All weekly dosage changes occurred on Sundays (no medication was administered on Saturday to allow for "washout") due to the inter-individual variation in serum and blood plasma levels following an acute administration of MPH (Gualtieri et al., 1982) and to control for potential rebound effects. Parents were instructed to give their child a medication capsule each morning, 1/2 hr prior to breakfast, and to tell them it was a vitamin supplement (i.e., to help control for expectancy effects, cf., Rapport et al., 1982). Both used and unused envelopes were returned to the CLC on a weekly basis to control for medication compliance. Weekly teacher ratings and daily classroom observations (described below) were completed during the morning hours due to the behavioral time-response course of MPH (Swanson, Kinsbourne, Roberts, & Zucker, 1978).

Teacher Ratings. Classroom teachers completed the ADD-H: Comprehensive Teacher Rating Scale (ACTeRS; Ullmann, Sleator, & Sprague, 1985) each Friday throughout the experimental conditions, which reflected the children's behavior during the morning hours (until 11:30 a.m.) of that week. Morning ratings were obtained due to the relatively short half-life of MPH (i.e., 2-3 hr). All teachers were blind as to when medication was administered and specific doses. The ACTeRS includes 24 items describing classroom behavior on four factors (attention, hyperactivity, social skills, and oppositional behavior), has adequate psychometric properties, and has been shown to be highly sensitive in detecting medication effects (Ullmann et al., 1985).

Behavioral Observations. Children were observed in their regular classrooms for 20 min intervals, 3 days per week across the 6-week evaluation period. Classroom observations began approximately 1 1/2 to 2 hr after the children's morning medication, during which time the class completed in-seat academic work assigned by the teacher. Children were unobtrusively observed by trained undergraduate and graduate students for 60 consecutive intervals during each observation period throughout the study. Each interval was divided into 15 s of observation followed by 5 s for recording. All observers were blind as to when medication was administered and specific doses.

A child's behavior was categorized as either on-task or off-task in a manner identical to that used by Rapport et al. (1985). Off-task behavior was defined as visual nonattention to one's materials for more than 2 consecutive s within each

15 s observation interval, unless the child was engaged in some other task-appropriate behavior (e.g., sharpening a pencil, asking the teacher a question).

Academic Measures. Children's performance on regularly assigned academic work during the scheduled observation periods was used as a dependent measure in an effort to preserve ecological validity, yet still maintain adequate experimental control. Academic seat-work typically involved completing arithmetic problems or language arts assignments. Assignments were graded after class by either the teacher or primary observer. Daily performance was recorded for both the percent of problems completed and the percent of problems completed correctly.

Reliability. An inter-observer reliability check of each child's on-task behavior was taken on 33% of the observation days, and at least once during each experimental phase. Overall reliability was consistently above 80%, with a mean of 88% across children. A mean Kappa value of .85 was obtained for all children.

Reliability checks of academic measures were completed daily for all children. Checks were made by either the primary observer in the classroom or the teacher (whomever did not initially score the child's paper for that day), and consisted of scoring the child's assignment independently of the other. Agreement was defined as agreement on the number of problems completed and number of problems completed correctly. Reliability was computed by dividing agreement between observer and teacher by agreement plus disagreement and multiplying by 100 to calculate the percentage. Observer-teacher agreement on problems completed and performance accuracy was consistently above 96% for all children across experimental conditions.

Results

A one-way multivariate analysis of variance (MANOVA) with repeated measures across dose (MPH) was performed using each of the six dependent variables (CPT omission and commission errors, ACTeRS attention factor scores, percent on-task, assignments complete, and assignments correct) and was highly significant ($F(30, 442) = 5.05, p < .001$) according to Wilk's lambda criterion. Follow-up univariate analyses of variance (ANOVA) with repeated measures across dose were performed separately for each dependent measure. Due to the number of planned follow-up tests, a Bonferoni critical value of .008 was established a priori to control for potential experimentwise error (i.e., univariate levels must exceed this value to be considered statistically significant).

Significant overall effects were found for CPT commission errors ($F(5,115) = 5.29, p < .001$), on-task behavior ($F(5, 115) = 23.64, p < .001$), percent of assignments completed ($F(5, 115) = 9.44, p < .001$), and ACTeRS attention factor scores ($F(5, 115) = 9.65, p < .001$). Percent of assignments completed correctly did not exceed Bonferoni critical values ($F(5, 115) = 3.19, p < .05$) and non-significant overall effects were found for CPT omission errors.

Newman-Keuls post hoc analyses were conducted for each of the significant univariate findings to examine between-dose differences on the various dependent measures. In the first of these analyses, the mean number of CPT commission errors (indicative of impulsive responding) under placebo was found to be significantly greater than those made under the 10 mg ($p < .05$), 15 mg and 20 mg ($p < .01$) conditions. A significantly greater number of commission errors were also made under 5 mg compared to the higher 15 mg ($p < .05$) and 20 mg ($p < .01$) doses. Similar between-dose differences were found for the length of time children spent attending to their academic assignments in the classroom. Children's on-task behavior was not significantly different during baseline and placebo conditions, but was significantly enhanced under the 10 mg, 15 mg, and 20 mg weekly MPH conditions ($p < .01$). Significantly higher on-task behavior was also found under the 10 mg ($p < .05$), 15 mg and 20 mg ($p < .01$) MPH conditions compared to 5 mg.

The mean number of weekly academic assignments completed by the children was significantly enhanced during the 10 mg, 15 mg, and 20 mg MPH conditions compared to baseline and placebo ($p < .01$), with the 10 mg ($p < .05$), 15 mg and 20 mg ($p < .01$) MPH doses also resulting in higher academic completion rates compared to the lower 5 mg dose.

A number of significant between-dose differences were found for the teacher ratings of children's classroom attention (ACTeRS attention factor). For example, children were rated as significantly more attentive under the 10 mg ($p < .05$), 15 mg and 20 mg ($p < .01$) weekly MPH conditions compared to baseline and placebo. They also received significantly higher ratings (indicative of improvement) under the 15 mg ($p < .05$) and 20 mg ($p < .01$) conditions compared to 5 mg, and under 15 mg and 20 mg ($p < .05$) compared to 10 mg. No further between-dose effects were found for any of the dependent measures.

In a subsequent analysis, tests for trend (Winer, 1969) were used to examine the shape of the relationship between dose (MPH) and each of the dependent measures. Although dependent variables showing non-significant univariate effects are not traditionally subjected to follow-up trend analysis, they were included herein (i.e., percent of assignments completed correctly and CPT omission errors) to determine whether a quadratic dose-response relationship (i.e., an inverted U-shaped function) may have precluded finding significant overall effects. A significant linear relationship between MPH dose and each of the dependent measures (percents on-task, assignments complete and correct, ACTeRS attention factor ratings, and CPT commission errors) was found except for CPT omission errors (see F values in Table 1), with all other potential relationships (e.g., quadratic, cubic) non-significant. Thus, the majority of dependent measures improved as a function of increasing dose in a linear fashion. A follow-up test for equality of slopes, with dose as the covariate and each dependent measure treated as a separate group, was non-significant (F (4, 590) = 1.31, N.S.), indicating that the significant linear relationships described above were of approximate equal slope and intercept.

The dose-response relationship between MPH and each of the six dependent measures was planned using both fixed-dose and milligram per kilogram (mg/kg) plotting methods. A comparison between the two methods was performed to address the question regarding whether response to MPH is dependent upon total body weight. The subsample of 6 children was randomly selected from the total N to facilitate graphical illustration and interpretation of individual dose responsiveness (i.e., drawn curves using all 24 children would be too difficult to visually interpret).

A series of t-tests was conducted to compare the dose-response relationship obtained for the subsample children (n = 6) with the total sample (N = 24) on each of the six dependent measures. No significant differences were found between the mean scores of the subsample and those of the total sample for each of these variables across dosages, indicating that the former were representative of the total group's response to MPH.

In examining the curves depicted in Figure 2 (Note: this Figure is presently being prepared), it appears that children's performance across the various dependent measures show similar dose-related changes using both fixed-dose and mg/kg plotting methods; thus lending support to the notion of individual responsivity to psychostimulant medication. Moreover, total body weight alone does not appear to adequately explain these differences, as several children with similar or identical weights (see mg/kg graphs in Figure 2) showed dramatically different response patterns under active drug conditions.

Although the response of most children to differing doses of MPH was consistent across tasks and behaviors in the present study, a few of the children evinced task-specificity. For example, two of the children showed continued gains or stability in percents on-task and assignments completed as a function of increasing dosage, however, both children's MFF test commission errors scores decreased at the higher dose levels.

Discussion

The results of the present investigation demonstrate a functional relationship between MPH and ADD-H children's attention in the classroom environment. As a group, children showed significant decreases in their CPT commission errors (indicative of reduced impulsivity) and corresponding increases in classroom attention (on-task behavior) and academic productivity across most MPH dosages compared to placebo. The dose-response relationship between the various dependent measures and increasing dosages of MPH was linear.

Curiously, the CPT's primary measure of attention, omission errors, was not significantly affected by any of the MPH dosages. Several explanations may have accounted for this result. For example, past investigators finding no significant drug effect on CPT omission errors reanalysed their data using only those children showing abnormal levels of omission errors during baseline conditions and found significant drug effects (Charles, Schain et al., 1979). To investigate this possibility, we conducted a similar post-hoc analysis using the 20 children showing abnormal baseline omission error scores (based on GDS standardization criteria), yet continued to find no drug effect. Given the significant dose-response effect of MPH on the other dependent measures, a more likely explanation of the results is that the GDS version of the CPT is not a medication sensitive measure for detecting changes in ADD-H children's attention (Swanson, Barlow, & Kinsbourne, 1979). Previous investigations have found similar results using both ADD-H children (Campbell, Douglas, & Morgenstern, 1971; Rapoport et al., 1980) and normal adults (Klorman, Bauer et al., 1984) treated with psychostimulants

The continued improvement in children's classroom attention and academic performance at higher dose levels was surprising, given that previous investigations have generally shown a deterioration of performance at doses exceeding 0.3 mg/kg (Brown & Sleator, 1979; Sprague & Sleator, 1977). At least two factors may have been responsible for these differences. The performance decrements at higher dosages reported in previous studies have largely occurred on laboratory-based tests of short-term memory (Sprague & Sleator, 1977) and impulsivity (Brown & Sleator, 1979), which may be more sensitive to drug effects. This explanation appears unlikely, however, as recent studies have shown increases in ADD-H children's performance on similar tests of short-term memory (Rapport et al., 1985), impulsivity (Rapport, Stoner, DuPaul et al., 1985), and attention (Rapport, DuPaul, Stoner, et al., in press) with dosages exceeding 0.3 mg/kg. A more likely explanation involves Sprague and Sleator's (1975) earlier hypothesis of continued cognitive improvement beyond 0.3 mg/kg, which could not be tested by the interpolated dose-response curve drawn between placebo, 0.3 mg/kg, and 1.0 mg/kg in the 1977 and 1979 studies. Thus, children may continue to evince improvement beyond 0.3 mg/kg, with performance decrements occurring as they approach 1.0 mg/kg.

In addition to the overall findings discussed above, a molecular analysis of individual children's response to MPH revealed several interesting results. In comparing the fixed-dose to mg/kg response curves for example, it was apparent that both methods of data plotting yield similar functional relationships between behavior and dose. Of particular interest was the idiosyncratic response children exhibited for a particular dependent measure (e.g., percent of academic assignments completed) despite identical or similar body weights.

Additional within-subject differences were apparent for a particular dose across behavioral domains. Although a clear majority of children showed consistent changes across the various dependent measures as a function of dose, the response of others to differing doses of MPH was task-specific (i.e., optimal responsivity depended not only on a child's reaction to a particular dose, but also on the behavior being assessed).

AUTHOR'S NOTE

The results presented herein are part of a larger clinical dose-response study presently being conducted by the Children's Learning Clinic to examine ADD-H children's CPT and academic performance. There will be approximately 36 children involved at the study's conclusion. Consequently, the results presented above may change, as would any interpretations offered to explain the results. We respectively request that no person cite the findings presented above without specific written permission of the authors. A complete manuscript of the study's results will be prepared upon completion of data collection. Appropriate graphs and figures will also be available at that time.

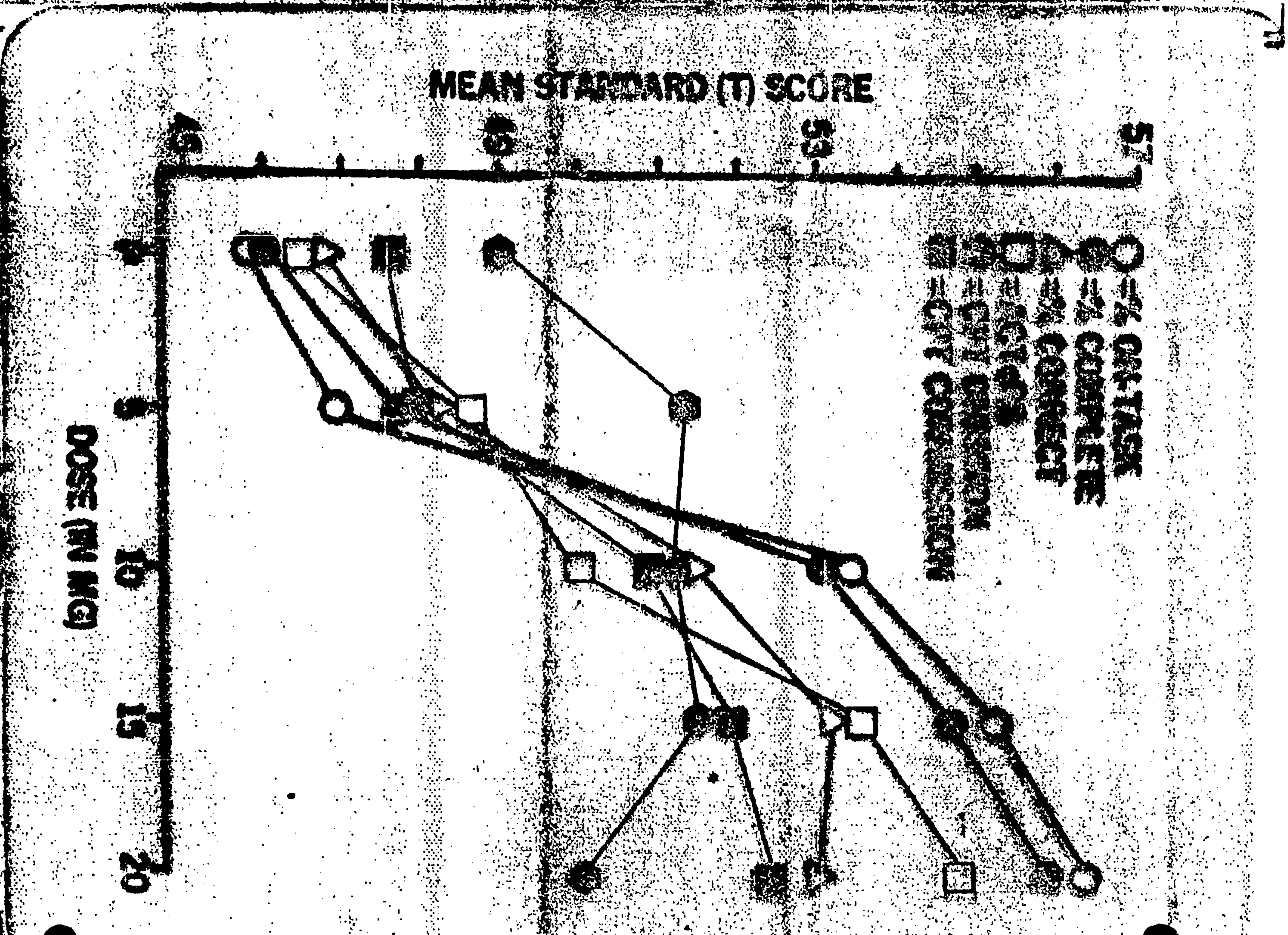
TESTS for TREND (with baseline omitted from analysis)

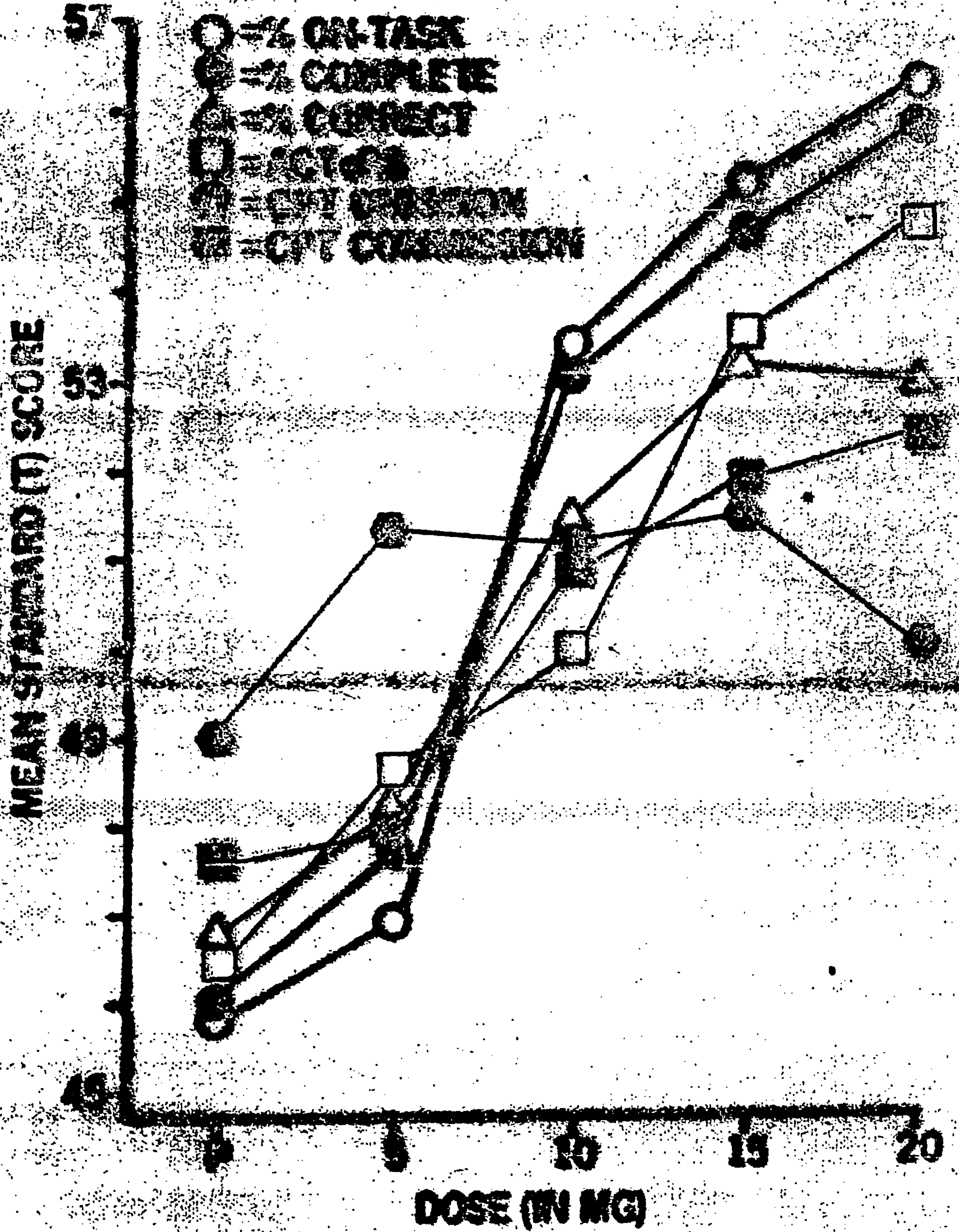
F Values for Trend Analysis

Measure	Linear	Quadratic	Cubic	Quartic	df
	FUNCTION				
% On-task	62.22***	1.14	2.74	2.06	1,92
% Complete	33.24***	.76	.74	.75	1,92
% Correct	13.79***	.81	.84	.09	1,92
ACTeRs	31.95***	.04	.24	.54	1,92
Omission Errors	.46	4.06	.06	.33	1,92
Commission Errors	29.43***	.39	1.46	.80	1,92

***$p < .001$

Test for Equality of Slopes (ANCOVA with dose as the covariate and each measure treated as a separate group):
$F(4,590) = 1.31$, N.S.


MEAN STANDARD (T) SCORE
57
DOSE (IN MG)
10
15
20


=% ON-TASK
=% COMPLETE
=% CORRECT
=CPT COMMISSION
57
53
49
45
MEAN STANDARD (T) SCORE
0
5
10
15
20
DOSE (IN MG)

Development of a Multi-Method Clinical Protocol for Assessing Stimulant Drug Response in ADD Children

Russell A. Barkley

Department of Psychiatry

University of Massachusetts Medical Center

and

Mariellen Fischer, Robert F. Newby, and Michael J. Breen

Department of Neurology

Medical College of Wisconsin

Running Head: Stimulants and ADD

Address for Reprints: Russell A. Barkley, Ph.D., Department of Psychiatry, University of Massachusetts Medical Center, 55 Lake Avenue North, Worcester, MA 01605.

Abstract

A protocol was developed for the clinical assessment of stimulant drug response in children with Attention Deficit Disorders (ADD) that is easy to administer. A triple-blind, placebo-controlled crossover design was used to assess two dose of Ritalin (0.3 mg/kg and 0.5 mg/kg BID) and a placebo. Each condition lasted one week. A one hour assessment battery was administered at the end of each drug condition. This consisted of parent and teacher ratings of ADD behaviors, problem settings, drug side-effects, laboratory tests of vigilance and impulse control, and clinic playroom observations of ADD behaviors during academic performance. Twenty-three of 28 children (ages 5 to 12 years) completed the entire study. Results demonstrated that most measures in the protocol were sensitive to both doses of medication. Eighty percent of the children were judged positive responders to medication. While minor revisions to the present protocol are suggested, it is clear that a quantitative, objective, cost-effective protocol for the clinical evaluation of stimulant drug response is feasible and should represent a substantial improvement over the subjective impressions now commonplace in clinical practice. Local centers equipped with the necessary resources for performing this type of objective evaluation should assist physicians by providing a standardized protocol for the early identification and rational pharmacological treatment of this significant mental health problem.

Development of A Multi-Method Clinical Protocol for Assessing Stimulant Drug Response in ADD Children

Stimulant drugs remain the most commonly recommended treatment for children with Attention Deficit Disorders (ADD), particularly in pediatric and child psychiatry clinics (Barkley, 1981, Dulcan, 1984). Yet, the decision to medicate such children is often based solely upon initial parental interviews and the titration of drugs solely on telephone conversations during the drug trial. While many decry the unreliability of such sources of information in general (Ullman, Egan, Fiedler, Jurenac, Pliske, Thompson, & Doherty, 1981) and point to the need for more systematic and objective evaluations of drug responding specifically (Cantwell, 1980; Ross & Ross, 1982), an objective, multi-method, yet cost-effective assessment protocol for use in clinical practice has remained elusive. Objective laboratory tests and direct observational methods for evaluating stimulant drug effects have been used for more than 15 years in the research literature (Barkley, 1977; Cantwell & Carlson, 1978) but are expensive, cumbersome, or unavailable to clinicians. However, recent developments in direct behavioral observation systems (Barkley, 1986a; Milich, Loney, & Landau, 1982) and micro-computer technology (Gordon, 1983; Klee & Garfinkel, 1983) allow for the adaptation of these previously cumbersome research instruments to the efficient clinical assessment of inattention, impulsivity, and other ADD behaviors in children.

This paper describes the development of a protocol for the clinical assessment of ADD children and their stimulant drug responses and its use with 28 clinic-referred ADD children. Varley and Trupin (1983) previously described a double-blind clinical protocol for evaluating stimulant drug responding but this procedure employed only brief parent and teacher ratings of behavior. Kinsbourne and Swanson (See Dalby, Kinsbourne, Swanson, &

Sobol, 1977) were probably the first to report the use of an objective instrument for clinically assessing stimulant drug responses, but they relied solely on one measure (paired-associate learning task) of only one type of behavior and typically conducted the evaluation of the child within a single day. More reliable and valid information can be obtained from assessments that rely on different sources of information (parent, teacher, clinician) from different settings (home, school, community), using different methods (ratings, laboratory tests, direct behavioral observations) combined into a single protocol (Mash & Terdal, 1981) which examines drug responding over a reasonable length of time. After reviewing the recent research literature, the present authors selected those methods demonstrating the most reliable sensitivity to stimulant drug effects while offering a minimum investment of time and resources for a clinical service.

The Drug Evaluation Protocol

Procedures

After receiving approval from the Human Research Review Committee of the Medical College of Wisconsin, the assessment service was advertised by letters sent to all pediatricians, child psychiatrists, and child neurologists in the greater Milwaukee area. Care was taken to explain the purpose and procedures of the service and to assure area clinicians that ongoing professional care of the ADD children would remain under their supervision. Clinicians wishing to refer children to the service provided a letter of referral, a separate letter authorizing school staff to administer the noon dose of medication (mandatory in Wisconsin public schools), and a prescription for both doses of methylphenidate (0.3 and 0.5 mg/kg given twice daily).

This prescription was sent to the pharmacy of an affiliated hospital for preparation of the two doses and placebo. The lactose placebo and appropriate dosages were prepared to the nearest 2.5 mg using regular Ritalin tablets (5 and 10 mg size) which were then crushed and placed within orange opaque gelatin capsules (size #o, Ely Lilly Co.). Capsules disguised both the differences in doses across the three drug conditions as well as the distinct taste differences between Ritalin and placebo. While each drug condition lasted 1 week, extra capsules were provided to permit the rescheduling of the clinic visits when necessary (e.g., child illness) without undue inconvenience in supplying families with more medication. Unused capsules were returned to the clinic staff at the end of each week to allow for a convenient check on compliance to the drug regimen.

Children were initially seen for a 1.5 hour evaluation consisting of the measures described below. This screening was conducted to insure that children referred for this service were appropriate for such a drug trial. Of 28 children between 5 and 12 years of age referred to the service during its first 9 months, 3 were deemed inappropriate for the drug trial based upon this initial evaluation. One child had a history of multiple tic disorder and was rejected because of possible increased tics with stimulants (Comings & Comings, 1984; Golden, 1982); a second child had a history of a cerebral vascular accident 1 year earlier and was rejected because of the known cardiovascular pressor effects associated with Ritalin (Cantwell & Carlson; Hastings & Barkley, 1978); and a third child was diagnosed as having only Oppositional Disorder of Childhood without evidence of ADD.

Those children deemed eligible for the protocol were then scheduled for 3 weekly evaluations and their parents provided with the first week's supply of the drug/placebo as well as the telephone numbers of the first author. If parents observed side effects of concern to them, they were instructed to contact the first author who would then notify the referring physician and discuss whether or not to discontinue the drug trial. Two of the 28 children initially referred were discontinued during the drug evaluation due to the development of tic reactions to the medication. Children were assigned to one of six possible drug orders of the placebo (P), low dose (L), and high dose (H) conditions: PLH, PHL, LPH, LHP, HLP, HPL.

The parents, children, their teachers, and the clinical assistant conducting the weekly drug evaluations were kept blind to the child's drug condition and order until the end of the entire evaluation. At that time, the drug code was broken and parents informed of the order and dose levels. The children were always evaluated on the last day of each drug condition and the clinical assistant insured that the children received a dose of Ritalin within 1 hour of the evaluation. After the final assessment, the findings of the entire evaluation were then reported to the referring physician and care of the child returned to his/her clinical practice. Recommendations for additional forms of treatment (e.g., parent training in behavior management, marriage counseling, individual therapy for parents evidencing depression, referral of child for special educational assessments, etc.) were occasionally made based upon the findings of the initial clinic evaluation of the children and their parents. Many of these therapies were also available at the clinic where the initial evaluations were conducted.

Assessment Methods

Initial Clinical Evaluation. The initial clinical evaluation consisted of the measures described below:

1. Semi-structured parental interview regarding: present behavioral, emotional, and learning problems; developmental, medical, and school histories; and family history and current family problems.

Ques K854903 84 3

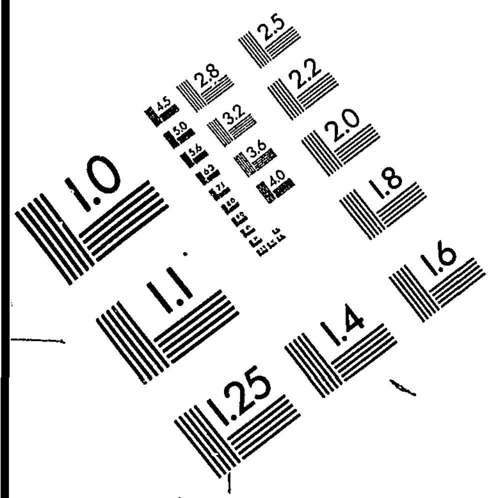
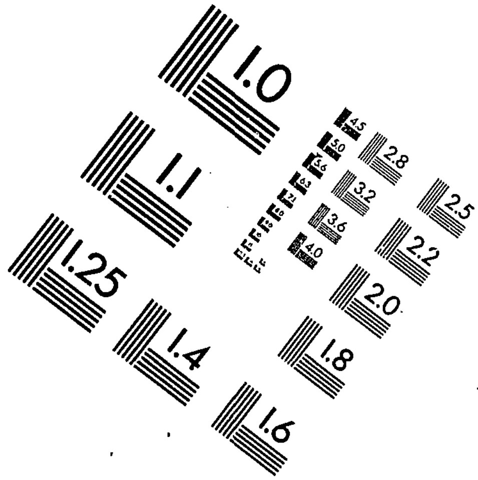
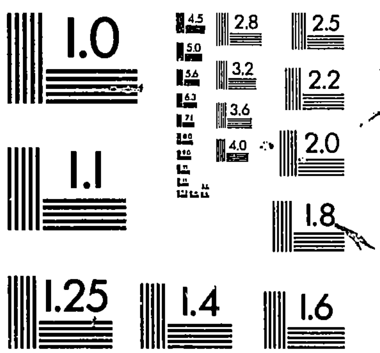
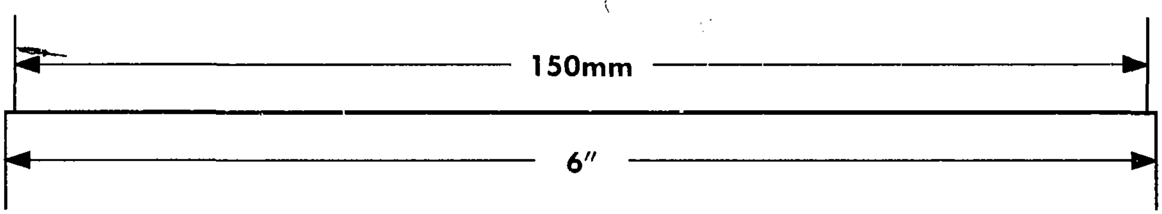
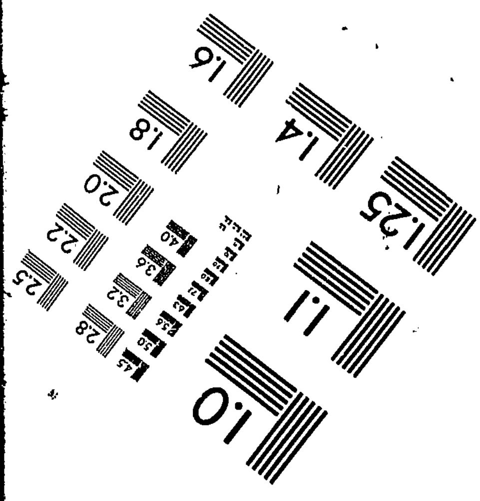


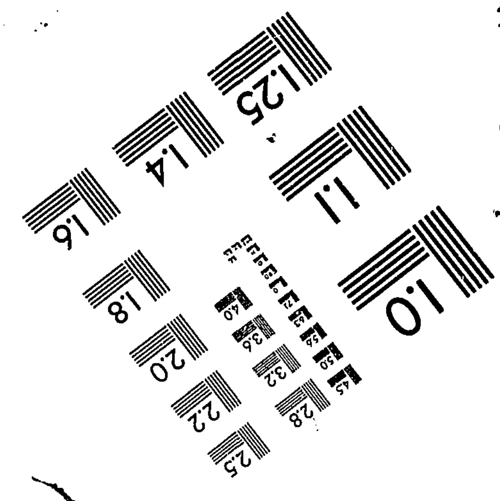
IMAGE EVALUATION TEST TARGET (MT-3)











PHOTOGRAPHIC SCIENCES CORPORATION
770 BASKET ROAD
P.O. BOX 338
WEBSTER, NEW YORK 14580
(716) 265-1600

K854903

2. Vigilance and Impulse Control: These constructs were assessed using the Gordon Diagnostic System (GDS: Gordon, 1983). This system used a small computer that permitted two tests to be administered to the child in a machine-paced procedure. The device was a metal box with a display screen on the front surface and a large blue button beneath it. Just above the screen were one red and one green light. The first test was a Delay Task believed to assess impulse control. It required that the child sit before the device, wait a short period of time, and then press the blue button. If the child waited a sufficient period of time (6 seconds) and then pressed the button, the green light appeared and the child earned a point. Cumulative points were displayed on the screen. If the child did not wait an adequate period of time before pressing the blue button, the green light did not appear and no point was awarded. The children were not told bu the experimenter how long to wait -- only that if they waited long enough they could earn a point. the task lasted 8 minutes and the child's scores were the number of correct responses (successful delays), the total number of responses, and the ratio of correct to total responses (Efficiency Ratio). Gordon (1979) has shown that the Delay Task significantly discriminates hyperactive from non-hyperactive clinic-referred children and is significantly correlated with teacher ratings from the Conners Teacher Rating Scale.

The second test was a measure of vigilance or sustained attention, similar to the continuous performance tasks used in stimulant drug research with ADD children (Barkley, 1977). The child sat before the device and a series of numbers were shown on the screen at a rate of one per second. The 6 to 12 year old children were to press the blue button whenever the 1-9 number sequence appeared. The 5 yaer old children were told to watch only for the number 1 to appear and then to depress the button. The task lasted 9 minutes. The child's scores were the number of correct responses, number of omissions (missed target stimuli), and number of commission errors. Normative data are available for children ages 3 to 16 years for both GDS tasks.

3. ADD Behaviors During A Restricted Academic Situation: Following the testing of the children, the experimenter took the mother and child to a playroom (12 x 18 ft.) equipped with a one-way mirror and intercom. The mother was asked to have her child perform a series of math problems. The child sat at a small table away from his/her mother. The mother was seated on a sofa situated across the room and read magazines. The mother was told to see to it that the child remained at the table, did not interrupt her, and completed the problems. three sets of math problems with varying difficulty levels were available. The problems were chosen from commonly available math workbooks at local educational stores. That set most appropriate for the child's age was used. Sets were available for ages 5-6, 7-8, and 9-11 years. Each set contained 5 pages of math problems to insure that the child did not finish the task before the 15 minutes of observation were completed. Three different versions of each difficulty level were used to reduce the likelihood of practice effects over the 3 weeks.

During this 15 minute period, the clinical assistant coded mother and child behavior. A tape recorder was used to cue the coder to the occurrence of every 30 second interval. During each interval, the coder scored the occurrence of each of eight behaviors. The category was checked as occurring only once regardless of the frequency of its occurrence during that 30 second interval. The categories were: Off-Task, Fidgets, Vocalizes, Talks to Mother, Plays with Objects, Out-of-Seat, Child Negative, and Mother Commands. A single score was derived for each category, this being the percentage occurrence of these behaviors relative to the total possible occurrences for that behavior. A total percent occurrence for all behaviors was also derived as a single Total ADD Behaviors score.

This observational system was selected to evaluate those disruptive behaviors most frequently seen in ADD children during situations many ADD children find most difficult; situations in which they are required to accomplish sustained tasks independently of adult assistance. Prior research has shown that a similar restricted academic situation and coding system not only discriminated ADD from normal children but also from non-ADD conduct problem children (Milich, Loney, & Landau, 1982). A similar coding system also showed significant stability over a 2 year period for normal and clinic-referred children (Milich, 1984; Milich, Loney, Whitten, 1983), and significant correlations with teacher ratings of hyperactivity and behavior problems.

The coder was trained to a level of 75 percent agreement with the first author using pre-recorded videotapes of mothers and children in this situation. Subsequently, intercoder reliability estimates were taken on five of the subjects by a second coder. Intercoder reliability was calculated using the number of agreements divided by the total possible number of occurrences. Inter-coder reliability averaged .77.

4. Parent Ratings of Child Behavior: While the child was being given the laboratory tasks, the mother completed several rating scales of her child's behavior. These were:

a. Child Behavior Checklist (CBCL: Achenbach & Edelbrock, 1983): This is a 118 item rating scale having two sections, one assessing Social Competence and the other Behavior Problems. The scale was used because it provides a generally comprehensive assessment of parent ratings of most of the commonly occurring dimensions of psychopathology in children and has been shown to discriminate ADD from normal children (Edelbrock, Costello, & Kessler, 1984). Normative data are available for ages 4 to 16 years for both sexes.

b. Conners Parent Rating Scale - Revised (CPRS-R; Goyette, Conners, & Ulrich, 1978): This is a 48 item questionnaire concerning various types of child behavior problems. The scale can be scored to yield five factors: Conduct Problems, Laerning Problems, Psychosomatic, Impulsive-Hyperactive, and Anxiety. The original version of this scale has been shown to be sensitive to stimulant drug effects (Barkley, 1977). While the revised scale has no research on its drug sensitivity, it was employed here because of its similarity to the original form while requiring far less time by the parents to complete. Although redundant with the information from the CBCL as part of the initial evaluation, the CPRS-R was used in this initial exam to reduce the likelihood that its repeated use in the 3 week drug protocol would result in practice effects that might confound its results. Such effects are known to occur with this instrument and primarily develop between the first and second administrations. It is recommended that it be administered once before using it to assess treatment effects (Barkley, 1988b).

c. Home Situations Questionnaire (HSQ: Barkley, 1981): This rating scale required parents to answer whether or not a child posed behavior problems in each of 16 situations listed on the scale (i.e., While Playing Alone, When Asked To Do Chores, In Public Places, etc.). If so, the parent indicated the severity of the problem using a scale of 1 (mild) to 9 (severe). Two scores were obtained, these being the total number of problematic settings and the mean severity score. The scale differentiates ADD from normal children (Barkley, 1981; Tarver-Behring, Barkley, & Karlsson,

1985) and is sensitive to stimulant drug effects (Barkley, Karlsson, Pollard, & Murphy, 1985).

5. Parent Self-Report Measures: While the children were completing the laboratory tests, mothers completed the following questionnaires. This was to provide a more comprehensive description of the children's family. Previous research has demonstrated that parenats of ADD children have significantly more problems with stress, depression, and marital discord than those of normal children (Befera & Barkley, 1985).

a. Parenting Stress Index (PSI: Burke & Abidin, 1980): The PSI is a 150 item multiple choice questionnaiore which yields six scores pertaining to child behavioral characteristics (e.g., distractibility, mood, etc.), eight scores pertaining to maternal characteristics (e.g., depression, sense of competence as a parent, etc.), and two scores pertaining to situational and life stress events. These scores can be summed to yield 3 domain or summary scores, these being Child Domain, Mother Domain, and Total Stress.

b. Beck Depression Inventory (Beck, 1967): This is a 21 item multiple choice questionnaire designed to permit a quick assessment of self-reported levels of depression in the respondent. The total score is the sum of the number of credits per item across all items. Higher scores reflect greater depression. The scale was used since prior research has shown that mothers of ADD children may be more depressed than those of normal children (Befera & Barkley, 1985).

c. Locke-Thomes Marital Adjustment Scale (Locke & Thomes, 1980): This is a brief 19 item multiple choice questionnaire designed to permit a quick assessment of marital discord. The score was the sum of the number of credits assigned to each answer across all items. Lower scores reflect greater marital dissatisfaction. Previous research found mothers of ADD children, particularly those of boys, to have higher ratings of discord than mothers of normal children (Befera & Barkley, 1985).

6. Teacher Ratings of Child Behavior: The children's primary teachers were mailed a packet of rating scales to complete and return. This packet included:

a. Child Behavior Checklist - Teacher Report Form (Achenbach & Edelbrock, 1983): This questionnaire has two sections, the first dealing with Adaptive Functioning at School and the second with Behavior Problems. The Adaptive Functioning section can be scored for five scales (School Performance, Working Hard, Behaving Appropriately, Learning, and Happy) or two summary scores, one for School Performance and the second for Total Adaptive Functioning (sum of remaining four scales). The Behavior Problem Scale has 113 items similarly worded to the Parent Report Form described above. Norms are available for two age groups (6 to 11 and 12 to 16 years) for each sex. The scales for 6 to 11 year old boys are: Anxious, Social-Withdrawal, Unpopular, Self-Destructive, Obsessive-Compulsive, Inattentive, Nervous-Overactive, and Aggressive.

b. Conners Teacher Rating Scale - Revised (CTRS-R; Goyette et al., 1978): The CTRS-R is a 28 item questionnaire constructed similarly to the CPRS-R described above. The factors are Conduct Problems, Hyperactivity, and Inattentive-Passive. The original version of this rating scale (Conners, 1969) from which the revised version was constructed is one of the most widely used and sensitive measures of drug responding in ADD children (Barkley, 1977). Again, like the CPRS-R, this scale is redundant with information from the Child Behavior Checklist - Teacher Report Form. However, because it has a significant practice effect between first and second administrations, it was administered as part of the initial evaluation to reduce the likelihood of practice effects on the scale confounding its results when used in the repeated drug assessment battery.

c. School Situations Questionnaire (SSQ: Barkley, 1981): This rating scale is similar to the HSQ except that the settings to be rated deal with situations in the school most likely to be problematic for ADD children. The scale lists 12 situations (i.e., During Individual Desk Work, In the Hallways, During Large Group Work, etc.) and the teacher was asked to indicate whether or not the child was a problem in each. If so, the teacher rated the problem on a scale from 1 (mild) to 9 (severe). Two scores were derived, these being the number of problem settings and their mean severity rating.

Repeated Drug Evaluation Battery. The vast majority of the measures described above were repeated each week during the drug evaluation. Those not repeated were the Child Behavior Checklist (both parent and teacher versions), and the parent self-report measures (PSI, Beck, and Locke-Thomes scales). These measures were not repeated because they have not been shown to be sensitive to stimulant drug effects and were not of interest in the 3 week drug assessment. They were used during the initial evaluation of each child as diagnostic aides or to evaluate various aspects of family functioning pertinent to determining future treatment recommendations for the child or family.

One measure not used in the initial evaluation but repeated during each week of the drug trial was a rating scale of side effects. The Side Effects Rating Scale has been used in previous drug studies to obtain information about the occurrence of 17 possible side effects known to occur with Ritalin (Barkley, 1981). Parents or teachers answered Yes or No as to whether each side effect was noted the previous week and, if so, they rated its severity using a scale from 1 (mild) to 9 (severe). Two scores were derived from both the parent and teacher completed scale: the number of side effects and their mean severity rating. These measures have been shown to be sensitive to dose effects of Ritalin (Barkley et al., 1985).

Results

Subjects

A total of 28 children between 5 and 12 years of age suspected of having ADD were referred to this drug assessment service. There were 21 boys and seven girls. All were believed to have ADD by their referring professionals. Of these 28, three were rejected for reasons described earlier, and two were prematurely discontinued during the evaluation due to tic reactions developing to the medication. The tics subsided within several days after discontinuing the medication in both children. Denckla, Bemporad, and MacKay (1976) found that fewer than 2 percent of a large number of ADD children receiving medication in their clinical practice developed tic reactions. Our results suggest that the reactions may be more common among ADD children exposed to stimulants than had previously been believed. Clinicians should question parents thoroughly about such movement disturbances during the drug trials.

To be eligible for the drug evaluation, the children had to meet the following criteria: (1) parent and/or teacher complaints of poor sustained attention, impulsivity, and restlessness; (2) a duration of at least 12 months of these behavior problems; (3) onset of the problems by 6 years of age: (4) scores on the Hyperactivity Factor of either the CPRS-R or CTRS-R of at least 1.5 standard deviations above the mean for same age normal children; (5) no history of mental retardation or significant developmental delay: (6) absence of epilepsy, gross brain damage, gross sensory deficits, or severe emotional disturbance (autism, psychosis, etc.); and (7) have no history of tic disorders or Tourette's Syndrome, cardiovascular problems, or previously poor response to Ritalin after age 5 years.

The initial characteristics of the 23 children (17 males, 6 females) completing the full evaluation and their parents are set forth in Table 1. These findings suggest that this sample was quite deviant from normal with

---

Place Table 1 about here

---

respect to parent and teacher ratings of behavior and pervasiveness of the behavior problems across situations. However, some children were identified as ADD (Hyperactivity Factor score greater than 1.5 standard deviations above the normal mean) by their teachers but not their parents on the Conners' scales while others were so identified by their parents but not their teachers. Many, however, were rated as ADD on these scales by both informants. The significantly deviant T scores on the Hyperactivity and Inattention factors of both the parent and teachers versions of the CBCL corroborate the success of the screening criteria in selecting a group of significantly inattentive, impulsive, and restless children. A majority of these children appeared to have difficulties with both ADD and conduct problems. Their mothers did not appear to be more disturbed or depressed than normal as indicated by the percentile scores on the PSI, but did report more behavioral problems and stress with their children on this measure. Similarly, the mothers did not rate themselves as more depressed than normal on the Beck Depression Inventory or as having more marital discord than normal on the Locke-Thomes scale, using commonly recommended cutoff scores for deviance on these scales.

On the lab measures, these children had a mean performance which fell in the borderline range (between 5th and 10th percentile) on the Delay Task and made more errors than normal on both comission and omission measures on the Vigilance Task of the GDS. Behavioral observations during the Restricted Academic Situation in the clinic playroom found these children to spend an average of 56 percent of their time engaged in off-task behavior and 35 percent in fidgeting; hardly surprising given the reasons for their referral but validating the success of this observational method at detecting these problems. The children engaged in vocal noises or self-directed speech 30 percent of the time, and spoke to their mothers against restrictions not to do so approximately 25 percent of the time. The children engaged in negative behaviors such as whining, refusal to do work, or tantrums approximately 7 percent of the time and their mothers initiated commands toward the children during 14 percent of the observations. Overall, ADD behaviors were noted during an average of 25 percent of the observation intervals.

Reliability of Measures

When utilizing measures in a repeated assessment battery such as this, it is essential that the measures have satisfactory test-retest reliability between repeated administrations where no treatment occurred in the interim. Several measures have such information available in the literature or from their developers. Gordon (personal communication) has reported test-retest coefficients for 90 normal children over a 30 to 45 day interval for the Delay Task of .77 for Total Responses, .68 for Number Correct, and .60 for Efficiency Ratio. For the Vigilance Task, coefficients were .6? for Total Correct, .72 for Comission Errors, and .80 for Omission Errors. For the CTRS-R, Edelbrock, Greenbaum, and Conover (1985) recently reported 1 week test-retest coefficients of .95 for Conduct Problems, .95 for Hyperactivity, .88 for Inattention, and .96 for the Hyperactivity Index. All coefficients were statistically significant in these studies.

In order to gain information about the test-retest reliability on the remaining measures, scores from the initial assessment were correlated with those taken during the placebo condition. Such a method will likely yield lower estimates of reliability than the more traditional method of using fixed time intervals between two administrations of a measure in that both the number of weeks (1 to 3) and hence the number of repeat administrations of the measures varied across subjects (e.g., some children received placebo the first week, others the second, and others the third week).

For the CPRS-R, coefficients across this 1 to 3 week interval were .65 for Conduct Problems, .57 for Impulsive-Hyperactive, and .53 for the Hyperactivity Index. For the HSQ, reliability was .66 for the Number of Problem Settings and .62 for the Mean Severity Rating. Coefficients for the same measures for the SSQ were .78 and .63, respectively. For the Total ADD Behaviors taken in the Restricted Academic Situation, reliability was .86. All coefficients were statistically significant. Considering that these are lower bound estimates of reliability, the actual test-retest reliability of these measures using more traditional procedures would certainly prove higher. Hence, the reliability of all measures seems quite adequate. No reliability was calculated for the Side Effects Rating Scale as it was not administered as part of the initial evaluation of the children but only during the repeated drug evaluations.

Drug Effects

All dependent measures used in the repeated assessment battery across the drug conditions were submitted to one-way (drug condition) analyses of variance with repeated measures. The means, standard deviations, and results from the statistical tests are set forth in Table 2. Significant main

Place Table 2 about here

effects for drug condition were noted on 16 of the 31 dependent measures. The majority of these effects were found on the teacher ratings and the behavioral observations taken in the clinic during the Restrictedd Academic Situation. Surprisingly, significant drug effects were not noted on the percent of math problems performed correctly during this playroom observation, but this may have resulted from the use of different types of math problems across the three testing occasions. The number and rated severity of side effects were also not significantly affected by these two doses of medication. In view of the many drug related improvements in behavior, this suggests that the dose levels were well within a judicious therapeutic range for these children.

Pairwise comparisons were conducted on the 16 measures having significant main effects and these results are shown in Table 3. On the vast majority of these measures, the low and

Place Table 3 about here

high doses were equally effective in improving the behavior of these children. Only on the behavioral category of "Plays with Objects" during the Restricted Academic Situation was a significant dose effect noted for the high dose compared to the low dose. While both doses resulted in a significant reduction in this behavior, the higher dose proved more effective. Although significant drug effects were noted on the Vigilance Task of the GDS, only the high dose of Ritalin resulted in significant improvements on both Comission and Omission Errors.

At the end of each child's evaluation, the results were reviewed by the psychologist supervising the drug trial and a recommendation was made to the referring physician as to which dose level, if any, appeared to have been most effective for this child. All results entered into this clinical decision with that dose making the greatest changes in behavior across the most measures and with the least side effects being chosen as the best dose for the child at that time. These clinical judgements resulted in 5 (20%) of the 25 children entering the protocol not recommended for any medication (2 of these were the children discontinuing prematurely due to tic reactions), 6 (24%) recommended for the low dose, 8 (32%) recommended for the higher dose, and 6 (24%) having a moderate dose between the high and low doses recommended. This latter recommendation was made where it was clear that both doses resulted in an effective drug response but the higher dose resulted in greater side effects for that particular child. Hence, a total of 80% of the children entering the protocol were recommended for at least some level of Ritalin following this drug-placebo evaluation. Comparisons on the initial evaluation measures of the responders and nonresponders were not undertaken because of the inadequate sample sizes per group to permit satisfactory statistical power in such comparisons.

Discussion

The results of this drug study suggest that an easily administered, cost-effective, multi-method assessment battery can be quite useful in the routine clinical evaluation of stimulant drug responding in ADD children. Such a test battery draws upon multiple sources and types of information using several methods of assessment and yields a wealth of clinically useful information on each case. The battery can be conducted in less than 1 hour per week, can be accomplished for a reasonable clinic fee (net cost for the 5 hours of evaluation for personnel, resources, drugs, and supplies was $120), and can be easily conducted by a paraprofessional with a short period of prior training in administering the tests (further enhancing its cost effectiveness).

The most useful measures for detecting stimulant drug effects were those derived from the teacher ratings and from the behavioral observations of academic performance in the clinic playroom. While some may be tempted to eliminate the use of the direct observational measures in favor of using only the more easily obtained teacher ratings, we recommend against it. First, the playroom measures proved equally if not more sensitive to the drug effects than did the teacher ratings of behavior in school. Second, it was our experience that such behavioral observations lend added credibility to the assessment protocol from the view of parents, teachers, and referring professionals in that they provided an objective assessment of the child's behavior apart from parent and teacher opinions. Third, when teacher ratings may not always be so readily obtained, in cases of poor teacher cooperation, during summer vacations, or when teacher ratings are believed to be unreliable, the clinic playroom measures are a reasonable substitute. Finally, it is useful in such assessments to build in a modicum of redundancy across measures in order to corroborate the drug effects noted on one method taken in one setting from one source against those from a different method, setting, and source.

Similarly, while it would seem at first glance more efficient to dispense with the parent ratings since they were of little value in detecting drug or dose effects, this, too, would be a mistake. These measures were useful in the initial diagnosis of the children and provided a necessary vehicle for parents to inform clinic staff of their observations across drug conditions. In some cases, these ratings revealed significant drug effects or side effects of use in further treatment planning. In fact, several parent ratings (hyperactive-impulsive behavior and mean severity score of problem behaviors) did show significant drug effects, particularly at the higher dose. Moreover, such ratings were valuable in providing ecologically valid assessments of behavior in an important setting (home) as reported by a very important observer (parents), whose views in part resulted in the initial referral of the child. Nor should it be at all surprising that drug effects were limited on the parent reported data as parents have much less opportunity to observe drug effects given their short time course (6 to 7 hours) and that such effects have often dissipated shortly after the child has returned home from school.

The failure of the GDS Delay Task to prove sensitive to drug effects was surprising in view of its purported evaluation of impulsivity -- a behavior typically responsive to stimulant drug effects -- and its previously demonstrated ability to differentiate ADD from non-ADD clinic-referred children (Gordon, 1979). Perhaps practice effects develop on this measure over repeated administrations (Douglas, personal communication) such that a ceiling effect develops beyond which medication effects cannot be detected. This would suggest that varying the delay interval across the three repeated assessments and forewarning the child in each session that the waiting time between button presses is different from that of the previous week might help to overcome this effect. The first author (RAB) is currently testing this procedural variation in an ongoing drug study. The Kagan Matching Familiar Figures Test (Kagan, 1966), a commonly used measure of impulsivity, is also being studied for comparison purposes because of its previously demonstrated sensitivity to drug effects (Barkley, 1977).

Significant drug effects were noted on the GDS Vigilance Task for both Comission and Omission Errors, but only at the high dose of Ritalin. This is difficult to explain considering that vigilance tasks of this sort are generally quite sensitive to lower doses of medication in testing periods even shorter than that used here (Sykes, Douglas, Weiss, & Minde, 1971; Werry & Aman, 1975). Perhaps extending the length of this test, say from 9 to 12 minutes, might enhance its sensitivity to the lower doses of medication. Nevertheless, retention of this measure in the drug protocol is recommended for the time being as it remains one of the few objective measures of inattention in the protocol aside from the direct observational measures and it has been shown to be dose sensitive in another drug study reported in this same journal issue.

The addition of a brief of verbal learning and memory task to this protocol might enhance the sensitivity of the laboratory tasks to stimulant drug effects while providing a measure of drug effects on a different domain of cognitive ability not presently assessed yet of import to academic performance. Numerous prior stimulant drug studies have found the paired associate learning task (PAL) to be sensitive to varying doses of stimulant medication (Dalby et al., 1977; Douglas, Barr, O'Neill, & Britton, in press; Rapport, Stoner, DuPaul, Birmingham, & Tucker, 1985). The particular version of this task used by Douglas et al. (in press) appears to be the easiest and least time consuming to administer, as well as the least expensive of these tasks.

The results of the present study argue strongly for the development of regional drug evaluation clinics to which area professionals can refer ADD children for objective, multi-method yet economical stimulant drug evaluations to replace the currently less reliable and more subjective approach employed in routine clinical practice. Many medical schools, hospitals, or mental health centers have the personnel, financial, and physical resources available to easily conduct these protocols and rapidly return useful information to the referring professional for the ongoing care and management of ADD children. Moreover, the use of a relatively consistent protocol such as this one across numerous regional drug assessment centers which combine their findings could also greatly accelerate the collection of scientific information on much larger subject samples than has heretofore been the case. In any event, the relatively high prevalence of the disorder and the widespread use of stimulant drugs to treat it demand no less than the most objective, reliable, valid, and cost-effective routine clinical assessment protocol our present scientific knowledge and technology can justify.

References

Achenbach, T. M., & Edelbrock, C. (1983). Manual for the Child Behavior Checklist and Revised Child Behavior Profile. Burlington, VT: Thomas Achenbach.

Barkley, R. A. (1977). A review of stimulant drug research with hyperacative children. Journal of Child Psychology and Psychiatry, 18, 137-165.

Barkley, R. A. (1981). Hyperactive children: a handbook for diagnosis and treatment. New York: Guilford Press.

Barkley, R. A. (1986a). Attention deficit disorders. In E. Mash & L. Terdal (Eds.) Behavioral assessment of childhood disorders (second edition). New York: Guilford.

Barkley, R. A. (1986b) A review of child behavior rating scales and checklists for research in child psychopathology. In M. Rutter (Ed.) Behavioral assessment methods for research in child psychopathology. New York: Guilford.

Barkley, R. A., Karlsson, J., Pollard, S., & Murphy, J. (1985). Developmental changes in the mother-child interactions of hyperactive boys: effects of two doses of Ritalin. Journal of Child Psychology and Psychiatry, 26, 705-715.

Beck, A. T. (1967). Depression: causes and treatment. Philadelphia, PA: University of Pennsylvania Press.

Befera, M., & Barkley, R. A. (1985). Hyperactive and normal girls and boys: mother-child interactions, parent psychiatric status, and child psychopathology. Journal of Child Psychology and Psychiatry, in press.

Stimulants and ADD

Burke, W. T., & Abidin, R. R. (1980). Parenting Stress Index (PSI): a family system assessment approach. In R. R. Abidin (Ed.) Parent education and intervention handbook. Springfield, IL: Charles C. Thomas.

Cantwell, D. (1980). A clinician's guide to the use of stimulant medication for psychiatric disorders of children. Developmental and Behavioral Pediatrics, 1, 133-140.

Cantwell, D., & Carlson, G. (1978). Stimulants. In J. Werry (Ed.) Pediatric psychopharmacology. New York: Bruner/Mazel.

Comings, D. E., & Comings, B. G. (1984). Tourette's Syndrome and Attention Deficit Disorder with Hyperactivity: Are they genetically related? Journal of the American Academy of Child Psychiatry, 23, 138-146.

Conners, C. K. (1969). A teacher rating scale for use in drug studies with children. American Journal of Psychiatry, 126, 152-156.

Daloy, J. T., Kinsbourne, M., Swanson, J. M., & Sobol, M. P. (1977). Hyperactive children's underuse of learning time: correction by stimulant treatment. Child Development, 48, 1448-1453.

Denckla, M. B., Bemporad, J. R., & MacKay, M. C. (1976). Tics following methylphenidate administration: a report of 20 cases. Journal of the American Medical Association, 235, 1349-1351.

Douglas, V. I. (1985). Personal communication, October, 1985.

Douglas, V. I., Barr, R. G., O'Neill, M. E., & Britton, B. G.

(1986). Short term effects of methylphenidate on the cognitive, learning, and academic performance of children with Attention Deficit Disorder. Journal of Child Psychology and Psychiatry, in press.

Dulcan, M. K. (1985). The psychopharmacologic treatment of children and adolescents with Attention Deficit Disorder. Psychiatric Annals, 15, 69-87.

Edelbrock, C., Costello, A., & Kessler, M. D. (1984). Empirical corroboration of Attention Deficit Disorder. Journal of the American Academy of Child Psychiatry, 23, 285-290.

Golden, G. S. (1982). Movement disorders in children. Developmental and Behavioral Pediatrics, 3, 209-216.

Gordon, M. (1979). The assessment of impulsivity and mediating behaviors in hyperactive and non-hyperactive children. Journal of Abnormal Child Psychology, 7, 317-326.

Gordon, M. (1983). The Gordon Diagnostic System. Boulder, CO: Clinical Diagnostic Systems.

Gordon, M. (1986). Personal communication, January, 1986.

Goyette, C. H., Conners, C. K., & Ulrich, R. F. (1978). Normative data for Revised Conners Parent and Teacher Rating Scales. Journal of Abnormal Child Psychology, 6, 221-236.

Hastings, J. E., & Barkley, R. A. (1978). A review of psychophysiological research with hyperkinetic children. Journal of Abnormal Child Psychology, 6, 413-447.

Kagan, J. (1966). Reflection-impulsivity: the generality and dynamics of conceptual tempo. Journal of Abnormal Psychology, 71, 17-24.

Stimulants and ADD

Klee, S. H., & Garfinkel, B. D. (1983). The computerized continuous performance task: a new measure of inattention. Journal of Abnormal Child Psychology, 11, 487-496.

Locke, H. J., & Thomes, M. M. (1980). The Locke Marital Adjustment Test: its validity, reliability, weighting procedure, and modification. Unpublished manuscript, University of Southern California.

Mash, E., & Terdal, L. (Eds.). (1981). Behavioral assessment of childhood disorders. New York: Guilford.

Milich, R. (1984). Cross-sectional and longitudinal observations of activity level and sustained attention in a normative sample. Journal of Abnormal Child Psychology, 12, 261-276.

Milich, R., Loney, J., & Landau, S. (1982). The independent dimensions of hyperactivity and aggression: a validation with playroom observation data. Journal of Abnormal Psychology, 91, 183-189.

Milich, R., Loney, J., & Whitten, P. (1983). Two year stability and validity of playroom observations of hyperactivity. Paper presented at the annual meeating of the American Psychological Association, Anaheim, CA.

Porteus, S. D. (1965). Porteus maze tests: fifty years of application. Palo Alto, CA: Pacific Books.

Rapport, M. D., Stoner, G., DuPaul, G. J., Birmingham, B. K., & Tucker, S. (1985). Methylphenidate in hyperactive children: differential effects of dose on academic, learning, and social behavior. Journal of Abnormal Child

Psychology, 13, 227-244.

Ross, D., & Ross, S. (1976). Hyperactivity. New York, Wiley.

Ross, D., & Ross, S. (1982) Hyperactivity (second edition). New York: Wiley.

Sykes, D. H., Douglas, V. I., Weiss, G., & Minde, K. K. (1971). Attention in hyperactive children and the effect of methylphenidate (Ritalin). Journal of Child Psychology and Psychiatry, 12, 129-139.

Tarver-Behring, S., Barkley, R., & Karlsson, J. (1985) The mother-child interactions of hyperactive boys and their normal siblings. American Journal of Orthopsychiatry, 55, 202-209.

Ullman, D. G., Egan, D., Fiedler, N., Jurenac, G., Pliske, R., Thompson, P., & Doherty, M. E. (1981). The many faces of hyperactivity: similarities and differences in diagnostic policies. Journal of Consulting and Clinical Psychology, 49, 694-704.

Varley, C. K., & Trupin, E. W. (1983). Double-blind assessment of stimulant medication for Attention Deficit Disorder: a model for clinical application. American Journal of Orthopsychiatry, 53, 542-547.

Werry, J. S., & Aman, M. (1975). Methylphenidate and haloperidol in children: effects on attention, memory, and activity. Archives of General Psychiatry, 32, 790-794.

Footnotes

1. Portions of this paper were presented at the meeting of the American Psychological Association, Los Angeles, CA, August, 1985. The authors wish to thank Kathy Schoendorf for her assistance with the data collection and transcription and Lewis Bloomingdale, M.D., for his comments on previous drafts of this manuscript.

Table 1

Initial Subject Characteristics

Measure	Mean	SD*	Range
Age (in years)	8.5	2.3	5.4 - 12.9
Child's Education (in years)	2.9	2.1	0 - 6
Age of Onset of ADD (in years)	4.6	1.5	1 - 6.9
Mother's Age (in years)	34.1	5.6	27 - 44
Mother's Education (in years)	13.3	1.9	11 - 18
Home Situations Questionnaire:			
Number of Problem Settings	9.1	4.0	2 - 16
Mean Severity of Problems	4.3	1.7	1 - 6.8
Conners Parent Rating Scale:			
Conduct Problems	14.7	7.3	3 - 27
Hyperactive-Impulsive	8.4	3.0	0 - 12
Hyperactivity Index	18.5	5.0	5 - 26
Child Behavior Checklist: (T scores)			
Hyperactivity	72.2	7.8	56 - 85
Delinquency	66.6	7.8	56 - 86
Aggression	71.2	10.8	55 - 88
School Situations Questionnaire:			
No. of Problems	7.4	2.8	2 - 12
Mean Severity	5.4	1.4	2.7 - 7.5
Conners Teacher Rating Scale:			
Conduct Problems	9.1	6.1	6 - 23
Hyperactivity	13.1	5.9	2 - 21

Stimulants and ADD

Inattention	14.4	5.5	6 - 24
Hyperactivity Index	17.9	5.4	8 - 26
Child Behavior Checklist - Teacher: (T scores)			
Inattention	70.2	8.4	56 - 87
Nervous-Overactive	69.1	7.9	55 - 86
Aggression	66.4	7.3	55 - 83
Beck Depression Scale	7.5	6.2	1 - 27
Locke-Thomas Marital Adjustment Scale	114.9	31.5	38 - 160
Parent Stress Index (Percentiles):			
Total Stress Domain	72.0	23.9	30 - 100
Child Domain	91.5	10.6	65 - 100
Mother Domain	54.8	31.1	10 - 95
Parent Depression Scale	54.1	29.6	10 - 95

* SD = standard deviation

Table 2

Means, Standard Deviations, and Statistical Test Results on the Dependent Measures for Each Drug Condition

Measure	Placebo Mean	Placebo SD	Low Dose Mean	Low Dose SD	High Dose Mean	High Dose SD	F	p
Home Situations Questionnaire:								
No. of Problems	8.0	4.5	7.4	4.2	7.3	5.1	0.54	-
Mean Severity	4.1	1.9	3.2	1.8	3.0	1.8	4.00	.05
Conners Parent Scale:								
Conduct Problem	10.7	7.7	8.8	6.4	9.3	8.0	0.87	-
Hyperactive	7.0	3.3	5.0	3.2	5.0	3.4	7.19	.01
Index	13.3	6.9	10.8	6.1	10.7	7.4	2.20	-
Side Effects - Parent:								
Number	5.0	4.0	5.1	3.9	5.2	3.3	0.08	-
Mean Severity	3.4	1.7	2.6	1.7	3.0	1.4	2.27	-
School Situations Questionnaire:								
No. of Problems	5.8	3.3	4.3	2.7	5.0	3.4	4.32	.05
Mean Severity	4.3	2.3	2.6	1.8	2.6	1.9	11.03	.01
Conners Teacher Scale:								
Conduct Problem	7.2	6.0	4.2	4.4	3.9	3.8	6.50	.01
Hyperactivity	11.1	6.5	6.7	5.1	6.8	5.4	7.97	.01
Inattention	10.8	5.0	8.0	4.7	7.6	5.1	7.60	.01
Index	14.3	7.1	8.9	5.7	8.7	5.4	10.69	.01
Side Effects - Teacher:								
Number	3.8	3.3	3.3	2.8	4.1	3.8	0.92	-
Mean Severity	3.0	2.3	2.6	2.1	2.9	2.3	0.34	-

Stimulants and ADD

GDS - Delay Task:								
No. of Rewards	40.1	13.0	41.4	19.0	41.7	15.2	0.19	-
No. of Responses	71.5	42.2	74.5	61.1	66.7	35.6	0.58	-
Efficiency Ratio	66.8	27.4	73.7	28.7	74.5	28.6	2.54	-
GDS - Vigilance:								
No. Correct	40.8	18.9	39.3	15.6	46.2	14.9	2.09	-
Commissions	14.5	19.4	16.7	22.4	6.1	6.2	3.41	.05
Omissions	10.0	12.2	9.6	15.5	4.5	7.1	3.24	.05
Playroom Observations (percent):								
Off-task	58.1	32.3	45.4	25.7	41.5	24.8	5.45	.01
Fidgets	38.3	25.4	27.4	20.8	30.4	23.3	2.10	-
Vocalizes	38.9	32.9	29.0	27.5	30.3	30.4	2.06	-
Talks to Mother	26.5	33.3	19.3	26.5	14.5	23.0	5.42	.01
Plays w/ Obj.	14.8	20.3	7.2	13.4	1.7	4.0	8.83	.01
Out-of-seat	12.6	24.7	5.4	16.4	2.5	8.7	5.82	.01
Child Negative	5.6	9.8	6.1	12.6	1.9	7.0	1.93	-
Mother Commands	15.8	25.6	9.7	19.4	7.2	13.9	3.35	.05
Total ADD Beh.	27.1	17.8	18.7	12.1	16.3	9.8	12.60	.01
Percent Correct in								
Math Problems	70.0	31.9	74.5	31.9	76.8	22.5	1.24	-

SD = standard deviation

F = F ratio from analysis of variance

p = probability value of F-test from analysis of variance

Table 3

Probability Values From Statistical Tests of Pairwise Comparisons on Measures Having Significant Main Effects

Measure	P vs L	P vs H	L vs H
Home Situations Questionnaire:			
Mean Severity	-	.01	-
Conners Parent Scale:			
Hyperactive/Impulsive	.01	.01	-
School Situations Questionnaire:			
No. of Problems	.05	-	-
Mean Severity	.01	.01	-
Conners Teacher Scale:			
Hyperactivity	.01	.01	-
Inattentive	.01	.01	-
Conduct Problems	.05	.01	-
Hyperactivity Index	.01	.01	-
GDS - Vigilance Task:			
No. of Commissions	-	.05	.05
No. of Omissions	-	.05	.07
Playroom Observations:			
Off-task	.05	.01	-
Talks to Mother	.05	.05	-
Plays with Objects	.05	.01	.05
Out-of-seat	.01	.05	-
Mother Commands	.05	.05	-
Total ADD Behaviors	.01	.01	-

Stimulants and ADD

P = placebo condition

L = low dose condition

H = high dose condition

FURTHER BACKGROUND ON VALIDATION OF GDS

Current GDS Research:

The Vicissitudes of Validation

Michael Gordon, Ph.D.

SUNY/Upstate Medical Center

Paper presented at a symposium entitled, "The Objective Assessment of ADD/Hyperactivity: Research on the Gordon Diagnostic System". Symposium presented at the 93rd Annual Convention of the American Psychological Association at Los Angles, California, August 1985.

My original intention for this presentation was to dazzle you with immense quantities of data about the Gordon Diagnostic System (GDS; Gordon, 1982). As it happens, we have just completed two major projects involving more than 800 subjects and a wide range of meaures. But in reviewing these data and considering time limitations, I felt that it would be worthwhile to err more on the side of selectivity than expansiveness in the material I presented. What I would like to do is use these data to g beyond the GDS, per se, and to illustrate what I see as several critical issues surrounding the evaluation of ADD/Hyperactivity.

All of us involved in the diagnosis and treatment of ADD/Hyperactivity are painfully aware of its complexities. This is a disorder comprised of multiple diagnostic dimensions, information about which is often based almost entirely upon the perceptions of different raters with varying internal norms, or upon objective measures which target differential aspects of functioning. But while we all acknowlege the complexity, most of us, myself very much included, tend to shy away from it. We talk more than we should about hyperactive children as if they were a relatively homogenous group which could be evaluated using agreed-upon criteria. We speak of one rating scale being better than the other in selecting the hyperactive child, or about which medications work and which do not, or about how this lab measure misses in the identification of hyperactive children while this one does not -- all with more enthusiasm than may be warranted.

Many of the complexities which need to be respected are illustrated in some of the data I will present to you now. If nothing else, they are intended to highlight how careful we need to be in approaching this diagnostic problem. They also dramatize how seemingly small changes in criteria for group definition can lead to significant changes in results.

What I plan to do is present to you the statistical trail of just one GDS-generated score -- that is, Total Commissions. My hope is that you will put yourself in the shoes of a test developer for a few minutes and confront the following problem: How do you best go about validating a diagnostic procedure for a disorder for which there are no satisfactory or agreed-upon benchmark criteria and where the entity being studied is multidimensional?

The Vigilance Task is a version ofthe venerable Continuous Performance Test (Rosvold et al., 1956). The child sits in front of the GDS and is told that numbers are going to flash on the screen one at a time. They are asked to press the button only when they see a "1" that comes right after a "9". The GDS records how many times they press the button correctly, how many hot "1/9" combinations they missed, and how many times they pressed the button when they weren't suppose to. This last score represents the Errors of Commission. Typically, they reflect the number of times, following appearance of a "1", the child pressed the button on a digit other than a "9". In other words, the child responds to any number following the "1" without waiting a few extra milliseconds to make sure it is a "9".

Now, one way to establish the utility of a score is to take what might be called the "Arrogant Face Validity" approach. Behind this strategy is the sentiment, "If you can't show me benchmark criteria against which I can establish the accuracy of my test, I'm going to go ahead and assume that the test does indeed measure a salient feature of the disorder. After all, children who don't attend well or who are impulsive when required to attend are bound to do poorly on this test. And besides, there's 25 years of research demonstrating the basic validity of this technique. So I am going to standardize it to the hilt and at least I'll be able to determine with some assurance the degree to which a particular score is normal or abnormal. If nothing else, I'll have some data based on a child's actual behavior".

This "Arrogant Face Validity" approach is one we have pursued intently over the past three years. We have gathered protocols on over 1000 nonhyperactive school children. It was important to me that we standardize the GDS on a nonhyperactive school sample as opposed to the more typical nonhyperactive clinic sample. As a clinician, I am more interested in how a child's performance compares to that of peers than I am in how they compare to the scores of children referred for a range of psychiatric problems. Parenthetically, the insistence that we gather data on a large nonpsychiatric population is one reason why the GDS is not microcomputer-based. To conduct large N studies, I needed an instrument that was light, portable, and absolutely reliable in its administration of the tasks. I also did not want to worry about the child being able to invade disk drives and adjustment buttons on monitors, or about variations in the pressures necessary to actuate a joystick.

Presented here are standardization data for the Total Commission score for children 3 - 16 years of age:

Standardization Table for Total Commissions

AGE	Normal	Borderline	Abnormal
6-7 (n=100)	0 - 10	11 - 23	≥ 24
8-9 (n=96)	0 - 8	9 - 21	≥ 22
10-11 (n=88)	0 - 3	4 - 15	≥ 16
12-16 (n=156)	0 - 3	4 - 11	≥ 12

This particular sample includes about 340 children. This is somewhat lower than the number we have for the Delay Task scores because that Vigilance Task came along a bit later in the development process. While we have means and standard deviations available, we have presented percentiles because they are more readily understandable to practitioners without much experience with statistics. Also, the distribution tends to be skewed to the negative side (that is, children tend not to have many errors of commission) so the percentile is a more accurate reflection of the distribution. Our practice has been to divide the scores into ranges. These include ABNORMAL which is the 5th percentile or lower, BORDERLINE which is the 6th through 25th percentile, and NORMAL, which is 25th percentile and above.

Now, it would be nice to leave it all at this table and say that we've done what we need to do. But, of course, most of us do feel the need to go beyond the "arrogant" approach to examination of cross-validation for this measure.

First let me review briefly some of the scales used in these studies. For information from parents, we have come to rely on the Child Behavior Checklist (CBCL; Achenbach, 1978) because of its fine standardization and large pool of items and factors. In some of our standardization studies we have used the Conners Parent Questionnaire (Goyette, Conners & Ulrich, 1978).

As for Teachers, the Teacher Rating Form (TRF; Edelbrock & Achenbach, 1980) is the companion version of the Child Behavior Checklist. You will hear most about the Innattentive and Nervous-overactive Factors as these are most related to issues of attention deficit.

Another fine measure has been developed by Ullman, Sprague and Sleator (1984) at the University of Illinois. The ACTeRs scale generates 4 factors, including Attention, Hyperactivity, Social Skills and Oppositional behavior. We also have used the Teachers Rating Scale (Goyette, et. al., 1978), scored both by the original and IOWA (Loney & Milich, 1982) formats, in some of the data you will be viewing.

Before dealing with issues of cross-validation and the commission score let's take a look for a moment at the nature of the criteria we're using to judge agreement:

Intracorrelations Among TRF Factors

(n=60)

	Nervous-Overactive	Aggressive	Social Withdrawal	Anxious	Internalizing
Inattention	.83	.64	.61	. 28	.50
Nervous-Overactive		.58	.45	.20	.34
Aggressive			.60	-.09	.46
Social-Withdrawn				.61	.83
Anxious					.93

The first point to be drawn from our data and, as you know, from the work of Loney and others, has to do with the high degree of overlap among factors that supposedly represent relatively independent dimensions of behavior. For example, there is a high degree of relationship among ratings of attention, aggression, oppositional behavior and even anxious behavior on the Teacher Rating Form.

Let me point out for this table as well as for the series you are about to see that the group used was comprised of 354 nonhyperactive subjects. I am showing you these data because it makes sense to see how parents and teachers in general rate children's behavior, before confounding the question with what happens when the children they are rating have been referred for psychiatric difficulties. Having said all that, the correlations among these scales, and the findings for the normal subjects very much hold when reanalyzed using clinical groups. While there is somewhat higher agreement in some areas, there is lower in others so the bottom line remains about the same.

You can see that the situation regarding overlap among factors is similar for the ACTeRS scale:

Intercorrelations Among Factors on the ACTeRS Scale

(n=198)

	Hyperactivity	SocialSkills	Oppositional
Attention	-.66	.73	.50
Hyperactivity		-.56	.63
Social Skills			-.56

Keep in mind that the Attention and Social Skills factors of the ACTeRS are scored in the opposite direction than the others. In other words, a high score on Attention indicates good attentional skills. Here are intercorrelations for Teacher's Conners:

Intercorrelations Among Standard & IOWA Conners' Factors

(n=291)

	I-O	Hyperactivity	Conduct Problem	Inattentive-Passive
<u>IOWA</u>				
Aggression	.76	.73	.65	.74
I - O		.93	.76	.77
<u>Conners</u>				
Hyperactivity			.68	.67
Conduct Problem				.43

The following table illustrates the same point but examines the intercorrelations among the various teacher rating scales:

ACTeRS, TRS, IOWA & TRF Intercorrelations

	ACTeRS		
	ATTENTION	HYPERACTIVITY	OPPOSITIONAL
TRS (n = 145)			
Conduct Problem	-.55	.56	-.83
Inattentive-Passive	-.81	.59	.43
Tense-Anxious	-.37	.22	.25
Hyperactive	-.68	.79	.70
IOWA (n = 145)			
I - O	-.77	.75	.72
Aggression	-.62	.59	.58
TRF (n = 51)			
Inattentive	-.87	.84	.47
Nervous-Overactive	-.73	.79	.47
Aggressive	-.46	.82	.89
Anxious	-.13	.18	.08
Internalizing	-.31	.37	-.28
Externalizing	-.73	.92	.75

Again there is a high degree of overlap among scales of hyperactivity, aggression, oppositional behavior, and even anxiety.

This table also demonstrates, though, that there is a fair degree of agreement between the various factors from teacher scale to teacher scale. Remember that these are ratings by one teacher of one child's behavior using two scales designed to assess the same behavior.

The situation regarding overlap is no better for parent ratings (Table 1). In fact, they seem to be even more intertwined. Most striking is the high correlation between the Hyperactivity Factor and Social Withdrawal.

These kinds of findings have led some to conclude that there are few differences between conduct problems and hyperactivity. An alternative explanation is that raters tend not to be very fine-grained in their judgments of behavior and often will lump together behavioral domains which, in reality, are distinct. A third hypothesis is that the rating scales don't ask the right questions and, consequently, do not allow the rater sufficient opportunity for differentiation among categories of behavior.

One obvious conclusion to be drawn, though, is that criteria such as teacher and parent rating scales are, by definition, going to include in the sample large numbers of aggressive and/or oppositional and also anxious children. In other words, study groups constituted using rating scales, rather than being homogeneous, will most likely represent various subgroups of hyperactivity or of other disorders which can be interpreted by a teacher or parent as hyperactivity.

The impact on us as test developers is clear: it is problematic verifying the sensitivity and specificity of a test when the "accepted" criteria themselves tend to be relatively non-specific and will form diverse, grab-bag groups of hyperactive subjects.

In this next table you see what we also already know -- parents and teachers tend not to agree very closely on whether or not a child is hyperactive, or the degree to which they exhibit these kinds of behaviors:

Essential Intercorrelations Between Parent & Teachers on Measures of Hyperactivity/Attention

TEACHERS		PARENTS CBCL - Hyperactivity
TRF		
	Inattentive	.47 (60)
	Nervous-Overactive	.40 (60)
TRS		
	Hyperactivity	.35 (265)
IOWA		
	I - O	.39 (264)
ACTeRS		
	Attention	-.43 (172)
	Hyperactivity	.31 (172)

Let me remind you that the TRF and CBCL scales contain almost identical items. The correlations between the CBCL Hyperactivity factor and the other scales related to hyperactivity also fall in a very modest range, particuarly when you consider that parent and teacher are rating the same child with instruments designed to assess the same domain of behavior. The situation appears worse for the two Conners scales. The correlations between the Parent and Teacher Hyperactivity Index was .25 (n=49).

Agreement between parent and teachers is particularly low when you ask them to rate aggressive and oppositional behavior:

Correlations Between Parent & Teachers on
Measures of Aggression/Opposition
TEACHERS

	TRS Conduct Problem	IOWA Aggression	ACTeRS Oppositional	TRF Aggression
PARENTS CBCL				
Aggression	.22 (265)	.22 (264)	.29 (173)	.24 (60)
Cruel	.20 (131)	.26 (131)	.29 (173)	.03 (25)

What is even more interesting is how the level of agreement between parents (CBCL) and teachers (ACTeRS) changes depending on the rating scale and the age of the child:

Intercorrelations Between Parent and
Teacher Ratings by Age

(AGE IN YEARS)	CBCL HYPER. & ACTeRS HYPER.	CBCL HYPER. & TRF INATT.	CBCL AGGR. & TRF C-P
6 - 7	.46 (36)	.37 (52)	.42 (52)
8 - 9	.47 (61)	.29 (87)	.19 (87)
10 - 11	.12 (42)	.39 (65)	.22 (65)

The correlation between the CBCL Parent Hyperactivity Factor and the Acter's Hyperactivity Factor, for example, ranges from .12 to .46 -- depending on the age of the child.

My intent in reviewing these data on the rating scales is not to deride them. Those of you familiar with our studies know that some of my best empirical friends are rating scales. I did want to underline a point made so often in the literature but sometimes forgotten. That is, that rating scales are far from being definitve criteria for constituting clinical groups. As you will see, you can select vastly different groups depending on what factors you employ and which raters or combination of raters you choose to rely upon. It is because of this state of affairs that I suppose I've come to bristle a bit at the question, "How accurate is the GDS" or the statement, "The GDS misses x percentage of children". Those statements imply that there is a definable target against which one could judge accuracy and that the pool of hyperactive children is a relatively homogenous one. It is also why, as a test developer, you constantly deal with being damned if you do and damned if you don't. When the GDS scores correlate with rating scales I will hear "That's not good because the rating scales are so inaccurate." When the correlations are low I hear, "Well, the GDS can't be very accurate because they don't correlate with the rating scales." The statisticians, in a way, resolve the entire issue by telling me that correlations are meaningless altogether.

With renewed respect for the issue of criteria confusion, we now move on to data regarding the validation of the Commission score. Presented in Figure 1 are median Total Commission scores for three groups: Children rated hyperactive by parents and teachers, children classified as non-hyperactive emotionally disturbed (ED), and nonhyperactive learning disabled (LD). The data are from dissertation studies by Peter Oppenheimer of University of Virginia and Deborah DiNiro of Syracuse University. Both the ED and LD groups are defined by placement. The LD children were all placed in special programs because they met New York State criteria for a learning disability. Almost all these children were reading-disabled. The ED group were all in psychiatric treatment at least once weekly. The bulk of the group were severely disturbed children who attended a day treatment program. You should keep in mind the level of severity of psychopathology in considering these data.

You will see in this figure that the Commission score differentiates clearly between these three major groups. Compared to the normal group, there is a dramatic difference in the median score -- from 13.5 for the hyperactive group to 4 for normals. These differences between the hyperactive group and the other two categories are very significant when analyzed statistically.

While things are looking good for the Commission score, it is only fair to present this next figure (see Figure 2). These are the various medians for hyperactive groups comprised according to a variety of criteria. That is, consensus hyperactives upon whom there is agreement between parent and teacher, hyperactives classified by parent or teacher, and those selected when you rely only on the parent or only on the teacher.

Regardless of how you comprise the groups, there is a difference between hyperactive and nonhyperactive. Nonetheless, you can also see that there are some very significant differences in median scores depending on what criteria you employ. Just look at the N for each group when you pick out "consensus hyperactives" as opposed to parent or teacher, or parent alone or teacher alone rated hyperactives.

We run into the same ambiguity when you ask the next logical question -- How many of the hyperactive group will be identified by the GDS as hyperactive?

Identification by Total Commission Score of Hyperactive Groups Constituted by Various Criteria

Rated ADD-H by:

	PARENT & TEACHER	PARENT OR TEACHER	PARENT ONLY	TEACHER ONLY	TEACHER ACTeRS
(AGE IN YEARS)					
6 - 7	71%	58%	62%	64%	--
8 - 9	42%	41%	46%	42%	-
10 - 11	72%	47%	45%	68%	--
TOTAL	61%	48%	51%	58%	68%
	(33/54)	(59/122)	(54/106)	(39/68)	(13/19)

Here you see rates of agreement when you select out from variously-constituted samples children who fall at the 10th percentile or lower on the different scales. (This level was used to allow for a sufficient N). We also used the 10th percentile for the GDS Commission score. You will note that the levels of agreement vary depending on the criteria for classification, the rating scale used, and the age of the child. Depending on your point of view, the GDS "misses" around 30% of hyperactive children or the parents and/or teachers overrated 30% of the sample as hyperactive.

To put this issue in perspective, I have here the same kind of exercise but this time we will look at the percentage of children picked out by the Acters Hyperactivity and Attention factors when applied to variously constituted groups of ADD/Hyperactive children:

Identification by ACTeRS Scales of Hyperactive Groups Constituted by Various Criteria

RATED ADD-H BY:

	PARENT AND TEACHER	PARENT OR TEACHER
ACTeRS		
Hyperactive	30%	18%
	(16/54)	(22/122)
Attention	54%	56%
	(29/54)	(68/122)

Remember the high correlations between the ACTeRS and the Achenbach scales and the fact that the Acters and TRF are filled out by the same teacher. You can see here that the levels of agreement are very low. In fact, they are generally lower than those achieved by the GDS.

The next concern regards the issue of false positives. How many nonhyperactive children will the commission score identify?

Identification of Nonhyperactives by
Total Commissions and ACTeRS

	SCHOOL NORMALS	COMBINED SCHOOL NORMALS & CLINIC NON-ADD
	(n = 50)	(n = 89)
TOTAL COMMISSIONS	12%	20%
ACTeRS		
Hyperactive	2%	6%
Attention	28%	34%

You will note that, at the 10% level, it selected 12% of a separate non hyperactive sample. You can see how that compares to the Acters scales. The clinic nonhyperactive sample includes the normal subjects plus those children rated by both parent and teacher as nonhyperactive. Out of 211 subjects, only 39 were rated by both as nonhyperactive. Incidentally, approximately 24% of the parents of 114 normal children in one of our studies responded that their children were overactive or impulsive.

So, at this point, despite all the caveats and complexities, you can see that this one score, Total Commissions, holds up reasonably well.

Now we move on to the correlations between Total Commissions and other lab measures:

Correlations Among Total Commissions, Fruit Distraction
Test, VADS & Other Child-Based Measures

	TOTAL COMMISSIONS	VADS TOTAL	FRUIT 4 - 1
FDT 4 - 1	.34	.41	
VADS TOTAL	.40		
TOTAL ER	.31	.10	.18
VTWD-CORR.	-.37	.57	-.30
VTWD-COMM.	.56	-.40	.33
VTAUD-CORR.	.18	.48	-.43
WRAT %	-.11	.41	-.26

The "FTD" stands for Fruit Distraction Test and was developed by Santostefano (1971). It is a kind of Stroop Color Distraction Test for children and requires them to name colors surrounded by various degrees of distractors and contradictions.

The Visual Aural Digit Span Test (VADS) was developed by Koppitz (1977) as a measure of intersensory integration and memory. The child is required to repeat digits presented either visually or aurally. The child's response is given either verbally or in written form. The Total Score represent an overall index of performance.

The next measures are correct and commission scores for two GDS enhancements we are developing. The Vigilance Task with Distracters (VTWD) is identical to the regular Vigilance Task except that random digits flash at random intervals on the outer two positions on the LED diplay. The Auditory Vigilance Task (VTAUD) is also identical to the standard Viglance Task except the digits are presented orally by means of a speech synthesizer.

Finally, the Wide Range Achievement Test (or WRAT) (Jastak & Jastak, 1978) generates an estimate of academic achievement.

The table indicates a modest degree of agreement among the various measures. Most of the correlations are statistically significant.

Conclusions:

Let me emphasize at the outset of my concluding remarks that my intent has not been to suggest that the GDS scores are better than rating scales or should supplant clinical judgment. Nor do I mean to imply that the situation is so arbitrary that despair and cynicism should abound. My main point is that, with such an important diagnosis, where there reigns so little consensus about diagnostic practice, and where there are criteria of such tenuous validity, it makes sense to throw out as wide a psychometric net as possible.

The inclusion of a laboratory measure, to me, appears essential to a sophisticated evaluation for ADD/Hyperactivity. If nothing else, it gives the clinician an opportunity to see for him or herself how a child performs in a standard situation which requires delay and sustained attention. Without behavior-based data, we are far more likely to overdiagnose this disorder, especially with boys. My hunch is that, more often than not, current practices tend to overlook girls with significant deficits in attention in the context of more manageable physical behavior

I support use of the GDS tasks because they are reliable, easy to use, well-standardized and well-supported by research efforts. Indeed, the feedback we get from practitioners is very positive and the validational studies are certainly encouraging.

As I mentioned at the outset, though, my point goes beyond issues regarding the GDS, per se.

These data re-emphasize what has been salient in the literature. The three major sources of data, that is information from parents, from teachers, and from the child him or herself via the lab measures tend to overlap only marginally. The child who emerges as abnormal in all three domains is going to be a rare, and probably quite difficult bird. While this consensus hyperactive has everybody agreeing, there is, of course, the questions of whether those other children who are considered hyperactive by two of the three realms could not benefit from treatment and might not represent important subgroups of the ADD/Hyperactive population.

These sort of data point to the need for us to become more intent on identifying subgroups of ADD/Hyperactive children beyond the "with and without hyperactivity" distinction. Delineations by "with and without aggression" make a great deal of sense but I am not sure we have good technology to make that distinction reliably. It was discouraging to us that the two scales of the IOWA Conners, that is the one presumably more related to inattention and the one related more to aggression, were so highly correlated.

I suppose one could become thoroughly cynical and suggest that the diagnostic categories should be ADD as rated by parent, ADD as rated by parent and teacher, etc.

These kinds of findings have convinced us all the more of the utility of the criteria offered by Dr. Barkley because they do throw out that wide net in an effort to establish severity, chronicity, and pervasiveness:

Diagnostic Criteria for ADD/Hyperactivity
(Adapted from Barkley, 1981)

1. Parent or teacher complaints of poor attention span, impulsivity, restlessness and inability to restrict behavior asa situation or adult demands.

2. Complaints of behavior place child 2 standard deviation (i.e. in 5th percentile from mean for his/her age and sex group as determined by well-standardized behavior scale of parent or teacher opinion.

3. BORDERLINE/ABNORMAL scores on either GDS task.

4. Onset prior to 5 years, 11 months (differs from DSM-III).

5. Duration of symptoms at least 1 year.

6. Pervasiveness of symptoms such that the child is scored at or below the 5th percentile for either the Home Situation Questionnaire (HSQ, 5th = 10) or School Situation Questionnaire (SSQ, 5th = 9).

7. IQ > 70.

8. Child does not display symptoms of autism or psychosis, or show evidence of blindness, deafness, aphasia or gross neurological disease (e.g. tumors, strokes, neurodegenerative disease, or obvious CNS trauma).

As you can see from this table, we would add BORDERLINE or ABNORMAL scores on the GDS tasks as additional criteria. We also have begun using different data for the Home Situations Questionnare and School Situations Questionnaire. Our data on 85 normal subjects suggest that the cuttoff for each should be 10 for the HSQ and 9 for the SSQ.

Finally, anyone concerned about the place of clinical judgment in this process should take great heart from the current situation. There is still plenty of room for sound judgment related to the integration of data and the appropriateness of conclusions. The complexity inherent in child behavior problems will, I am sure, guarantee the value of clinical skills for a long time to come.

References

Achenbach, T.M. (1978) The child behavior profile: I. Boys aged 6-11. Journal of Consulting and Clinical Psychology, 46, 478-488.

Barkley, R.A. (1981) Hyperactive children: A handbook for diagnosis and treatment . New York: The Guildford Press.

Edelbrock, C. & Achenbach, T.M. (1984). The teacher version of the Child Behavior Profile I: Boys aged 6 - 11. Journal of Consulting and Clinical Psychology , 52,(2), 20744-217.

Gordon, M. (1982). The Gordon diagnostic system, Littleton, Colorado: Clinical Diagnostics, Inc.

Goyette, C.H., Conners, C.K. & Ulrich, R.F. (1978). Normative data on revised Conners parent and teacher rating scales. Journal of Abnormal Child Psychology, 6, 221-236.

Jastak, J.F. & Jastak, S. (1978). Wide range achievement test: 1978 revised edition. Wilmington: Jastak Associates.

Koppitz, E.M. (1977). The Visual Aural Digit Span Test. New York: Grune & Stratton.

Loney, J. & Milich, R. (1982). Hyperactivity, inattention, and aggression in clinical practice. Advances in Developmental and Behavioral Pediatrics, 3, 113-147.

Rosvold, H., Mirsky, A., Sarason, I., Bransone, E. & Beck, A. (1956). A continuous performance test of brain damage. Journal of Consulting Psychology, 20, 343-352.

Santostefano, S. (1971). Fruit distraction test: A procedure for assessing in children the cognitive principle of "field articulation". Boston University School of Medicine, Boston, Mass: Author.

Ullmann, R.K., Sleator, E.K. & Sprague, R.L. (1984). A new rating scale for diagnosis and monitoring of ADD children. Psychopharmacology Bulletin, 20,(1), 160-164.

TABLE I

INTRACORRELATIONS AMONG CBCL FACTORS
(N = 345)

	AGGRESSION	HYPERACTIVE	CRUEL	ANXIOUS	SOCIAL WITH.	INTERNAL	EXTERNAL
AGGRESSION		.75	.72	.42	.60	.77	.97
HYPERACTIVE			.56	.44	.60	.75	.86
CRUEL				.40	.43	.52	.72
ANXIOUS					.41	.65	.45
SOCIAL WITHDRAWAL						.76	.64
INTERNAL							.81

FIGURE I

Median Total Commission Score by Diagnostic Category

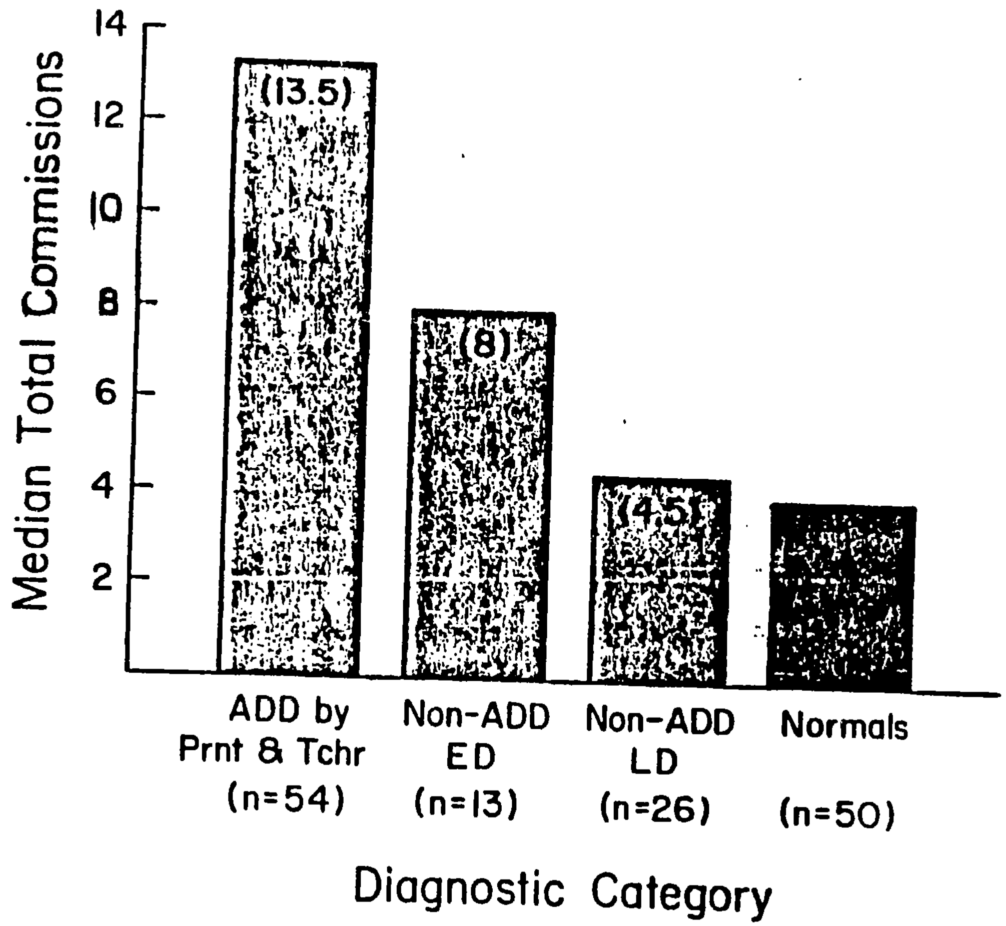
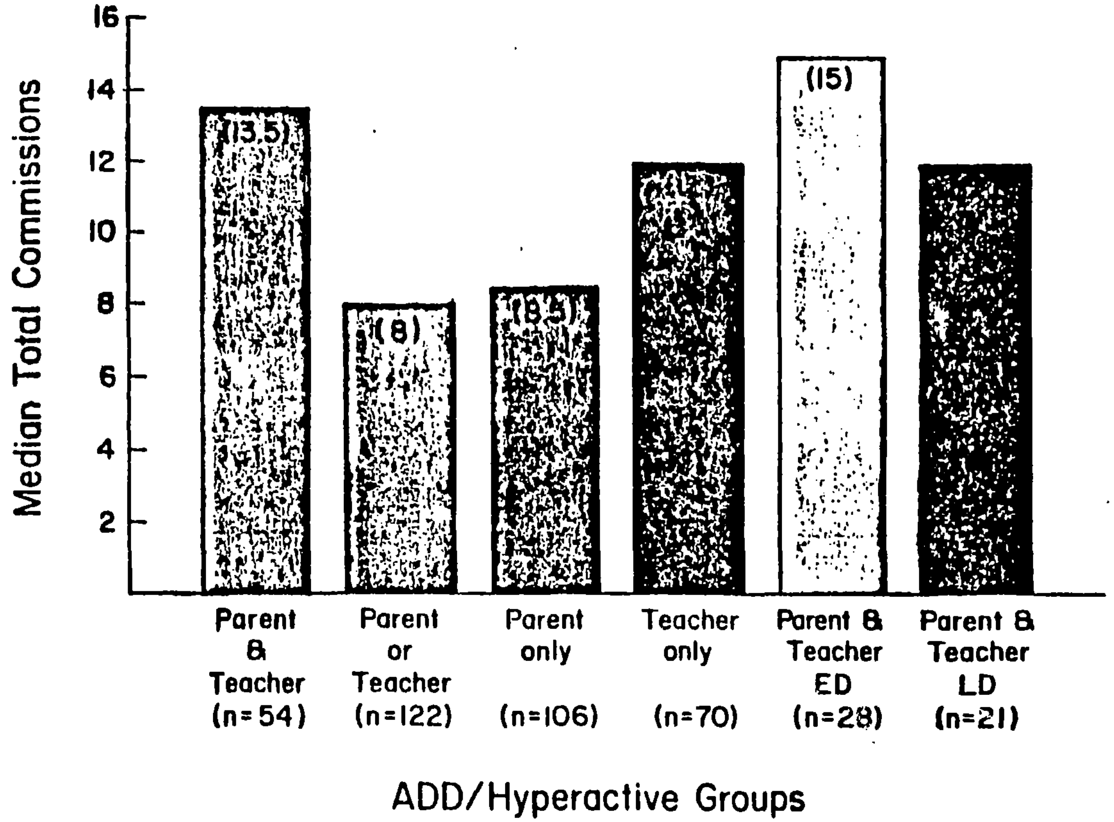


FIGURE II

Median Total Commission Scores for ADD/Hyperactive Groups Formed by Various Criteria



RS232C communications port. Software is available which allows for the direct transmission of GDS data to a microcomputer where it can be tabulated, graphed, stored, and compared to normative data. Connection of the GDS to a microcomputer enables collection of some ancillary data which cannot be extracted from the stand-alone unit.

The portability and ease-of-operation of the GDS has allowed for large-scale data collection. The standardization sample is comprised of 900 boys and girls from 3 to 16 years of age (Gordon & Mettelman, 1985). An additional 1,100 hyperactive and nonhyperactive protocols from various subject populations, including the deaf, blind, Spanish-speaking, emotionally disturbed, and learning disabled, have also been gathered.

A series of validational studies has shown that these game-like tasks differentiated accurately between hyperactive and nonhyperactive children from both outpatient, and day treatment settings (Gordon, 1979; McClure & Gordon, 1984). In a sample of school-referred children, the GDS distinguished children with ADD from those classified as reading disabled, overanxious, and normal (Gordon & McClure, 1983). Over 20 university and medical center research sites are currently conducting investigations involving the GDS. Most studies concern the effectiveness of the GDS for determining drug responsivity and for evaluating the success of pharmacotherapy (Atkinson, 1985; Barkley, 1985; Rapport, 1985). Others are investigating the use of the GDS for evaluating children with known brain damage, preschoolers considered "at-risk" for impulsive behavior, and the relationship between GDS scores and observational measures of classroom behavior. The GDS is also being used by school systems and pediatric practices across the country for the clinical evaluation of ADD-H.

Comments and Conclusions

A chronic impediment in the field of ADD-H has been the paucity of universal criteria and procedures for subject selection and treatment monitoring. Even in the realm of behavioral assessment, where truly creative approaches to testing children have been developed, procedures have generally not been sufficiently practical or well-standardized to allow for acceptance by the majority of professionals (Loney, 1981). As such, most continue to rely on rating scales or tests, such as the Wechsler Intelligence Scale for Children (WISC-R) (Wechsler, 1974), which have been shown to be of limited usefulness in differentiating between groups (Douglas, 1972). Further- ance of reliable administration and valid interpretation. The growing acceptance of the GDS within the professional community also permits greater comparison of data across studies.

It must be emphasized, however, that the GDS was never intended as a divining rod for ADD-H. This is a complex disorder which represents an array of subgroups and interactive diagnostic dimensions. The GDS is seen as an important tool to be used only in conjuction with other selection criteria and clinical judgment.

References

Atkinson, A.W., Cohen, P.C., and Kelly, P.C. Attention deficit disorder: The effects of ritalin on self-esteem a comparison of ACTeRS teacher scale, Conners' parent scale, and Gordon Diagnostic System in diagnosis and management. Paper presented at the American Academy of Pediatrics Meeting, Atlanta, Georgia, April 1985.

Barkley, R.A. Assessment of stimulant drug responding in ADD-H children. Paper presented at the 93rd Annual Convention of the American Psychological Association, Los Angeles, California, August 1985.

Conners, C.K. *Continuous Performance Test* [Computer program]. Kensington, Maryland: Behavioral Medicine, Inc., 1980.

Davenport, W. Vigilance and arousal: Effects of different types of background stimulation. *J. Psychol, 82*, 339–346, 1972.

Douglas, V.I. Stop, look and listen: The problem of sustained attention and impulse control in hyperactive and normal children. *Can. J. Behav. Sci.*, 4:259–282, 1972.

Doyle, R.B., Anderson, R.P., and Halcomb, C.G. Attention deficits and the effects of visual distraction. *J. Learn. Disab.*, 9:59–65, 1976.

Gordon, M. The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive children. *J. Abnorm. Psychol.*, 7:317–326, 1979.

Gordon, M., and McClure, F.D. The objective assessment of attention deficit disorder. Paper presented at the 91st Annual Convention of the American Psychological Association, Anaheim, California, 1983.

Gordon, M., and McClure, F.D. Assessment of attention deficit disorders using the Gordon Diagnostic System. Paper presented at the 92nd Annual Convention of the American Psychological Association, Toronto, Canada, 1984.

Gordon, M., and Mettelman, B.B. *Threshold Tables for the Gordon Diagnostic System*. (Available from Clinical Diagnostics, Inc., 300 E. Mineral Avenue, Suite 6, Littleton, Colorado 80122): 1985.

Hiscock, M., Kinsbourne, M., Caplan, B., and Swanson, J.M. Auditory attention in hyperactive children: Effect of stimulant medication on dichotic listening performance. *J. Abnorm. Psychol.*, 88:27–32, 1979.

Loney, J. Evaluating treatments for childhood hyperactivity: Some methodological considerations. In K.D. Gadow, K.D., and S. Loney, J. (eds.), *Psychosocial Aspects of Drug Treatment for Hyperactivity*. Boulder, Colo.: Westview Press, 1981, pp. 77–103.

Klee, S.H., and Garfinkel, B.D. The computerized continuous performance task: A new measure of inattention. *J. Abnorm. Child Psychol.* [illegible] 1983

allow for acceptance by the majority of professionals (Loney, 1981). As such, most continue to rely on rating scales or tests, such as the Wechsler Intelligence Scale for Children (WISC-R) (Wechsler, 1974), which have been shown to be of limited usefulness in differentiating between groups (Douglas, 1972). Furthermore, comparison of studies across research programs has often been hampered by idiosyncratic measures or sets of criteria.

The GDS represents an effort to establish a standard procedure for evaluating certain aspects of attention and self-control. While it is more costly and, in certain respects, less flexible than software-driven approaches, its portability, ruggedness, and extensive base of normative data offer the user greater assur-

Loney, J. Evaluating treatments for childhood hyperactivity: Some methodological considerations. In K.D. Gadow, K.D., and S. Loney, J. (eds.), *Psychosocial Aspects of Drug Treatment for Hyperactivity*. Boulder, Colo.: Westview Press, 1981, pp. 77–103.

Klee, S.H., and Garfinkel, B.D. The computerized-continuous performance task: A new measure of inattention. *J. Abnorm. Child Psychol.*, 11(4):489–496, 1983.

Margolis, J.S. *Academic Correlates of Sustained Attention*. (Unpublished thesis, University of California, Los Angeles, 1972).

McClure, F.D., and Gordon, M. The performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. *J. Abnorm. Child Psychol.*, 12(4):561–572, 1984.

Rapport, M.D. Comparing classroom and clinic measures of ADD: Dose response effects. Paper presented at the 93rd Annual Convention of the American Psychological Association, Los Angeles, California, August 1985.

Rosvold, H.E., Mirsky, A.F., Sarason, I., A continuous performance of brain damage. *J. Consult. Psychol.*, 20:343–350, 1956.

Weschsler, D. *Weschsler Intelligence Scale for Children-Revised*. New York: Psychological Corporation, 1974.

Microprocessor-Based Assessment of Attention Deficit Disorders

Michael Gordon, Ph.D.[1]

Dissatisfaction with formulating a diagnosis of ADD with hyperactivity (ADD-H) based almost entirely upon clinical judgment or the perception of raters has spawned the development of behavior-based assessment procedures. These efforts typically involve administration of the Continuous Performance Test (Rosvold, et al., 1956), or related measures of attention and self-control (Davenport, 1972; Doyle et al., 1976; Hiscock et al., 1979; Margolis, 1972). Although bulky and expensive electromechanical devices had previously been required, researchers are now programming microcomputers to administer these laboratory tasks to children (Conners, 1980; Klee & Garfinkel, 1983).

Use of the microcomputer has certain advantages, including flexibility of administration, the ability to store multiple data points, and, if a microcomputer is already available, cost-effectiveness. This approach, however, does have drawbacks for the researcher and, in particular, the clinician. The size of even portable microcomputers can make testing in multiple sites cumbersome. It is perhaps for this reason that standardization samples for software-driven programs tend to be, at best, limited in number and breadth. The transport of disk drives, monitors, keyboards, and the computer, itself, from location to location often discourages use of the procedure in other than a single research or clinic setting.

Another potential limitation of microcomputer-based testing stems from concerns surrounding the reliability of administration. Computer monitors vary in the size, shape, intensity, and color of displayed characters. Each brand of computer presents to the child a different keyboard, and accepts different types of peripherals, such as "joysticks, which are often used as manipulanda. Without control over such critical factors as the amount

ministration of the task. Efforts to secure the microcomputer are often unsatisfactory because they usually involve additional hardware, cabling, and expense.

An alternative approach to testing attention and self-control has been developed by the author (Gordon & McClure, 1983, 1984). The goal of this project has been to establish a practical, reliable, and well-standardized procedure available to both researchers and clinicians. The Gordon Diagnostic System (GDS) is a microprocessor-based, portable unit which, without an external microcomputer, allows for the administration of multiple tasks. The Delay Task, based upon a DRL operant schedule requires the child to inhibit responding in order to gain a reward. Specifically, the child is instructed to press a button, wait, and then press the button again. If she or he refrains from responding for at least 6 seconds, a light flashes and a reward counter increments. If the child responds before the interval elapses then the timer resets and no reward points are recorded. The GDS Delay Task generates three major scores: the number of responses (button presses), the number of correct responses (i.e., Correct) and the Efficiency Ratio which represents the percentage of correct responses.

The GDS also contains the Vigilance Task, which is a version of the Continuous Performance Test (Rosvold et al., 1956). The child is presented with a series of digits on the front display and is told to press the button every time a "9" appears that had been preceded by a "1". The GDS records the number of correct presses, the number of times the child fails to press the button when the "1/9" combination appears (i.e., errors of omission), and the number of extraneous button presses (i.e., errors of commission). For the testing of younger children, the GDS contains a "1" mode, which requires the subject to press the button every time a "1" appears. The same performance measures are recorded.

Although normative data were gathered using standard settings for task parameters, the design allows the user to select a wide range of parameters. This feature enabled the modification of parameters for the testing of adults as well as very young children.

Because the GDS is controlled by a programmable microprocessor, additional tasks can be explored. One such enhancement being studied is a version of the Vigilance Task which assesses the impact of distractors on a child's ability to sustain attention. This Distractibility Task is identical to

...ing the reliability of administration ... young children.

Computer monitors vary in the size, shape, intensity, and color of displayed characters. Each brand of computer presents to the child a different keyboard, and accepts different types of peripherals, such as "joysticks, which are often used as manipulanda. Without control over such critical factors as the amount of pressure necessary to actuate a switch, or the accuracy of timed sequences, repeatable administration across multiple microcomputers cannot be assured.

Finally, ADD-H children, despite a reputation for academic underachievement, routinely display a fine facility for disassembling delicate equipment. Unless the examiner is immediately present, the highly impulsive child will inevitably stick a finger in disk drives, unplug cables, adjust monitors, or in some other fashion interrupt standard ad-

Because the GDS is controlled by a programmable microprocessor, additional tasks can be explored. One such enhancement being studied is a version of the Vigilance Task which assesses the impact of distraction on a child's ability to sustain attention. This Distractibility Task is identical to the standard Vigilance Task except that random digits flash at random intervals on the outer positions of the LED display. The subject is still required to press the blue button when a "9" comes right after a "1". The only difference is that numbers flash on either side of the center (i.e., relevant) digit. Other enhancements under development are designed to evaluate attention across sensory modalities.

Even though the GDS administers tasks independently of a microcomputer, it can communicate with external hardware via an

[1]State University of New York, Syracuse, NY 13210.

*Pharmacology Bulletin (in press) Vol. 22 (1).

CHAPTER IN J.LONEY (ED.).
THE YOUNG HYPERACTIVE CHILD
HAWORTH PRESS (IN PRESS)

HOW IS A COMPUTERIZED ATTENTION TEST USED IN THE DIAGNOSIS OF ATTENTION DEFICIT DISORDER?

Michael Gordon, Ph.D.

SUNY/Upstate Medical Center

The advent of computerized assessment of behaviors associated with Attention Deficit Disorders (ADD) represents another step forward in the effort to define reliable and meaningful diagnostic criteria. These techniques were born of concern about the extensive degree to which diagnostic decisions were founded upon opinion, or upon data from traditional psychological tests of limited relevance to issues surrounding this disorder. Although clinical judgment, behavior rating scales, and clinical interviews are critical to a sophisticated evaluation, each approach harbors well-documented limitations. In the face of considerable evidence to the contrary, one would be hard put to justify the statement "I know an ADD child when I see one" or the practice of using only rating scale scores or IQ factors to formulate a diagnosis.

While not intended as magic geiger counters for ADD, computer-based laboratory measures do offer both the researcher and clinican an opportunity to incorporate data derived from a child's actual behavior. Unlike the other clinical methods, laboratory tasks generate objective data about a child's ability to perform in situations tailored to strike at the characteristic weaknesses of an ADD child. As such, they generally require a child to sustain attention and control behavior over a period of time and with varying degrees of feedback regarding performance. The tasks are often variants of attentional measures, such as the Continuous Performance Test (Rosvold, Mirsky, Sarason, Bransome & Beck, 1956), in which a child must respond only to a specific combination of symbols in a stream of irrelevant symbols. Many years of research indicate that ADD children fare poorly when required to attend and to develop a strategy for self-control.

Laboratory measures have had a long history in ADD research and a very short one in actual clinical practice. Research versions have traditionally been bulky, expensive, and impractical. There has also existed a certain degree of resistance to using mechanical measures within an evaluation which traditionally has relied heavily on the clinician's judgment. The burgeoning growth of computer technology has solved the technological limitations of earlier electromechanical devices. The increasing popularity of computers in clinical practice has also led to greater comfort with computerized assessment of behavior.

Considerations involved in selecting a computerized assessment technique are essentially identical to those important for the evaluation of any psychological test. The computerization of a measure does not obviate the need for evidence concerning reliability of administration, test-retest reliability, robust standardization, or meaningful studies of validity. Most of the available techniques have enjoyed only limited research investigation of psychometric properties. These procedures have generally been standardized on small groups of subjects, usually children who have been referred to a psychiatric clinic but for reasons other than those associated with ADD. While such samples may suffice for certain research purposes, the clinician is interested more in how a particular child's performance compares to that of peers than to the scores of children referred for a range of psychiatric problems.

Extensive standardization of computerized assessment procedures are particularly critical in the light of the legendary variability inherent in the test performances of children. It is also important to keep in mind that most of the scores generated by these measures are age-related. Without a very substantial normative base, misdiagnosis can result from failure to adjust for a child's developmental status.

Another key consideration concerns the practicality of an assessment procedure. If a technique is not designed to fit comfortably into the daily practice of a busy clinican, it will likely fall from use regardless of its diagnostic efficacy. While software-driven assessment programs have been available for many years, they generally have not found their way into clinical practice, often because clinicians have considered them too cumbersome to administer and score. Most practitioners have neither the time nor the computer expertise to struggle with complicated instructions or lengthy procedures. Most of the techniques also fail to meet the clinican's need for portability. Since the majority of those involved in serving children referred for ADD travel among schools and/or offices, they are loathe to carry around bulky equipment or to expend effort in connecting cables and attaching peripherals.

The clinician would, therefore, be wise to select a procedure that can be employed with ease and confidence. Along these lines, the assessment program should be accompanied by clear operating instructions, support materials regarding interpretation, technical data, and active service and research support.

Unfortunately, there are only a handful of available computerized techniques which even begin to approach the criteria mentioned above. Most rely on microcomputers to generate the tasks (Conners, 1980; Klee & Garfinkel, 1983; Greenberg, 1985) in contrast to a self-contained, microprocessor-based unit dedicated to task administration (Gordon, 1982). The advantage of the former strategy is that these software programs tend to be less expensive, more flexible in the varieties of tasks that can be administered, and more amenable to data storage. However, they also tend to be less transportable, rugged, practical, and reliable in administration. As indicated above, software packages tend to be poorly normed if normed at all. The one program that has enjoyed some standardization is the Garfinkel Assessment Battery, but this technique is restricted to research purposes.

The most widely-used procedure is the Gordon Diagnostic System (GDS; Gordon, 1982), developed by the author. Although supported extensively by research efforts, the GDS was designed specifically for clinical use. The GDS is a microprocessor-based, portable unit which, without an external microcomputer, allows for the administration of multiple tasks. The Vigilance Task requires the child to inhibit responding under conditions that make demands for sustained attention. A series of digits flashes one at a time on an electronic display. The child is told to press the button every time a "1" is followed by a "9". The GDS records the number of correct reponses, the number of time the child failed to respond to the "1/9" combination, (i.e. Errors of Omission), and the number of extraneous button presses (i.e. Errors of Commission). For the testing of younger children, the GDS contains a "1" mode, which requires the subject to press the button every time a "1" appears. The same performance measures are recorded.

The Delay Task requires the child to inhibit responding in order to earn points. Specifically, the child is instructed to press a button, wait a while, and then press the button again. If s/he refrains from responding for at least 6 seconds, a light flashes and a reward counter increments. If the child responds before the interval lapses then the timer resets and no reward points are recorded. The Delay Task yields three primary scores: the number of responses (button presses), the number of correct responses (i.e. Correct) and the Efficiency Ratio, which represents the percentage of correct responses.

The administration of both tasks takes less than twenty minutes. Although normative data were gathered using standard task parameters, the design allows the practitioner to select a wide range of settings. This feature enabled the modification of parameters for the testing of adults as well as very young children. The internal microprocessor generates the tasks and records quantitative features of a child's performance both for the entire session as well as for individual time blocks. In this way, the pattern of a child's performance across the session can be analyzed.

Because the GDS is controlled by a programmable microprocessor, additional tasks can be explored. One such enhancement being studied is a version of the Vigilance Task which assesses the impact of distraction on a child's ability to sustain attention. This Distractibility Task is identical to the standard Vigilance Task except that random digits flash at random intervals on the outer positions of the electronic display. The subject is still required to press the blue button when a "9" comes right after a "1". The only difference is that numbers flash on either side of the center (i.e. relevant) digit. Other enhancements under development are designed to evaluate attention across sensory modalities.

The portability and ease-of-operation of the GDS has allowed for large-scale data collection. The standardization sample is comprised of 1200 boys and girls from 3 to 16 years of age (Gordon & Mettelman, 1985). An additional 1100 hyperactive and nonhyperactive protocols from various subject populations, including the deaf, blind, emotionally disturbed, learning disabled, and Spanish-speaking have also been gathered. A series of validational studies has shown that these game-like tasks differentiated accurately between hyperactive and nonhyperactive children from both outpatient and day treatment settings (Gordon, 1979; McClure & Gordon, 1984). In a sample of school-referred children, the GDS distinguished children with ADD from those classified as reading disabled, overanxious, and normal (Gordon & McClure, 1983). Over 20 university and medical center research sites are currently conducting investigations involving the GDS. Most studies concern the effectiveness of the GDS for determining drug responsivity and for evaluating the success of pharmacotherapy (Atkinson, 1985; Barkley, 1985; Rapport, 1985). Others are investigating the use of the GDS for evaluating children with known brain damage, preschoolers considered "at-risk" for impulsive behavior, and the relationship between GDS scores and observational measures of classroom behavior. The GDS is also being used by school systems, mental health professionals, and pediatricians across the country for the clinical evaluation of ADD/Hyperactivity and assessment of therapeutic outcome.

Feedback from clinicians indicates that GDS testing tends to be conducted early in the diagnostic phase in an effort to achieve a more efficient evaluative process. The rationale for this approach is that GDS results can help direct the course of the ensuing evaluation by initially screening for general levels of self-control and attentiveness. For those children who perform within normal limits on the GDS and brief rating scale measures of hyperactivity, further assessment can be geared toward examining other possible explanations for the child's behavior aside from ADD/Hyperactivity, per se. Conversely, the clinician is likely to pursue more intently consideration of a diagnosis of ADD/Hyperactivity for those children whose GDS scores fall in the abnormal ranges.

Practitioners also report that the GDS testing to be most useful in helping to rule out a diagnosis of ADD/Hyper-activity for the many children who are not hyperactive and would best benefit from treatments other than those applied for ADD/Hyperactivity. The procedure is also used extensively in monitoring the effectiveness of pharmacotherapy. Finally, clinicians have consistently reported that, quantitative scores aside, the opportunity to observe a child perform in situations demanding of attention and self-control has been valuable.

Clinical examples:

Presented below are a series of cases in which GDS testing was included within an evaluation (adapted from Gordon, 1984,1985). While a few case histories, truncated in their description because of space limitations, cannot convey the full impact of objective testing, they can serve to illustrate the sorts of instances where computerized assessment can be meaningful.

Case 1

A nine year old boy was referred with complaints of restlessness, noncompliance, fighting with other children, and poor academic performance. While his parents had been recently divorced, these behaviors were longstanding and had been problematic since early childhood. A medical examination by his pediatrician was unremarkable, and he was referred for a psychological evaluation.

Upon initial contact with the psychologist, this youngster appeared inhibited, quiet, and withdrawn, giving no indications of impulsive or hyperactive behavior. However, when confronted with the demands of the Delay Task, the boy was unable to maintain the facade. He achieved an Efficiency Ratio of .44, which falls in the Abnormal Range. His behavior throughout the Delay Task was disjointed, involving much out-of-seat activity and extreme restlessness. While these behaviors in some ways helped him to suppress responding, because he was otherwise occupied, he was nonetheless unable to refrain from emitting a large number of unreinforced responses. This pattern of behavior was repeated during the Vigilance Task. He consistently responded to the digit immediately after the appearance of a "1", without waiting to see if it was a "9". His performance was suggestive of an inability to delay, particularly once he had been primed to respond. Following a complete diagnostic evaluation the youngster was classified as having ADD with Hyperactivity and was placed on a moderate dose of stimulant medication. His academic programming was geared more toward accuracy than speed, and he received resource help to encourage him to modulate his response style. Follow-up contact indicated substantial improvement.

CASE #2

Kevin, a 10-year old boy, was referred for a psychological evaluation by both his school and pediatrician. According to all concerned, Kevin was consistently impulsive, disruptive, inattentive, hyperactive and underachieving. Judging from rating scales and case-history information, his behavior met all the DSM-III criteria and the more stringent set of markers proposed by Barkley. The age of onset for his problems was before 5 years 11 months, his symptoms were chronic and pervasive, his IQ was well over 70, and ratings of his behavior by teachers on standardized checklists were beyond the 5th percentile. His school had demonstrated ample patience with his misbehavior to the extent that they allowed him to leave the classroom and run the hallways for five minutes when he could no longer contain his energy. At the time of the referral, the school was reaching the end of their capacity to deal with Kevin.

Kevin was administered the Delay Task, Standard Vigilance Task, and the Vigilance Task with Distractors. The Delay Task Efficiency Ratio of .75 was in the Borderline Range for his age. The most noteworthy feature of his performance was the very significant degree of variability across the four Time Blocks. At times he inhibited perfectly while at other times his performance approached the Abnormal Range. This marked inconsistency in the context of a profile which shows evidence of the capacity for delay is pattern found often in children whose problems have a strong emotional component. Unlike the more typical ADD profile in which there is little if any evidence of successful inhibition, the emotionally disturbed child's protocol will tend to show clear swings across the session between adequate and inadequate coping.

Kevin's performance on the standard Vigilance Task also fell in the Borderline Range (albeit the near normal end) for both Correct and Commissions. Again, there were indications of inconsistency of performance in that the middle Time Block, unlike the other two, was normal. As is so often the case, Kevin's behavior during the session was telling. Upon missing the "1/9" combination, he would slap himself in the face and complain that the numbers flashed too quickly for him. The examiner noted that Kevin became so anxious about performing on the task that he had trouble concentrating. His discomfort increased on the Distractibility version of the Vigilance Task in which his performance in all respects was in the Borderline Range.

The GDS evaluation indicated that Kevin, at times, could be more impulsive and inattentive than age-mates. However, his poor control did not consistently reach an abnormal range. In fact, he demonstrated a capacity to delay and attend adequately but had difficulty maintaining good performance for reasons that seemed, at least in part, related to emotional issues. His behavior during the testing was of a boy who became unusually anxious and self-denigrating when he met up with frustration or failure.

On the basis of this kind of GDS protocol, a complete psychodiagnostic evaluation is typically suggested. In Kevin's case, a full battery of psychologic tests was administered (WISC-R, WRAT, PIAT, Figure Drawings, TAT, Rorschach, Bender. etc.), in addition to clinical interviews with the boy, his family and teachers. The overall conclusion of this extensive evaluation was that Kevin's problems of self-control were secondary to a severe and chronic emotional disturbance. The onset of his hyperactivity at age 5 coincided with the gunshot murder of his mother. Because his biological father was unavailable (Kevin was born out-of-wedlock), Kevin was raised by his maternal grandparents. The grandparents were very devoted to the boy but, because of their own fears, suspicions and ambivalence about their custody of Kevin, they seriously limited his contacts with others. They also engaged in what appeared to be bizarre rituals around the memory of the boy's mother. This family history as well as a host of other factors left Kevin with intense fears and a dependent, hostile relationship with his grandparents. As he grew older he found it increasingly difficult to manage his anxieties and anger. He would become so flooded by stress that he often would have difficulty organizing himself for even simple tasks. Intensive psychotherapy both with Kevin and his grandparents was initiated.

CASE 3

A seven year old girl was referred to a child development clinic by her teacher for poor school performance, not following instructions, missing assignments, and difficulty with reading. The teacher viewed much of this youngster's behavior as willful and oppositional, and was at her wit's end as to how to help the girl learn. Her parents reiterated the teachers complaints, and added that they felt the child to be distant and aloof much of the time. During the clinical interview this youngster was pleasant and cooperative, but gave the impression of either wanting to be somewhere else or of simply daydreaming.

On the Delay Task she obtained an overall Efficiency Ratio of .82, well within the Normal range. She earned 46 Total rewards. Her responses appeared controlled, orderly, and goal directed. Obviously her difficulties did not lie in the area of impulse control. On the Vigilance Task her difficulties became quite clear. Although she produced just one error of Commission, which is well within normal limits, she made 33 errors of Omission, scoring well beyond the mean on this measure.

Although her motivation throughout the task was quite good, she was unable to maintain the degree of preparedness required to perform effectively on this part of the task. Therefore, we were able to identify her main difficulty as a deficit in sustained attention in the absence of impulsivity, which fits the DSM III classification of Attention Deficit Disorder without Hyperactivity.

Following a more extensive evaluation, she was placed on a very small dose of stimulant medication, and her teacher was informed of the results of the testing. Intervention strategies were offered for both home and school, and her adjustment improved in a satisfactory manner.

Case 4

James, a 13 year-old boy of average intelligence, was referred for GDS testing from an adolescent inpatient unit where he was undergoing treatment for a range of psychiatric problems. The precipitant of this second admission to the unit was the boy's attempt to choke himself with a pencil. James had a history of firesetting, suicidal ideation, poor peer relationships, harming animals, and academic underachievement. Full psychological evaluations had pointed to severe pathology across domains of functioning. He was described as anxious, impulsive, aggressive, depresssed and distractable. In addition to conduct problems, James was considered hyperactive and, upon occassion, had received brief trials of stimulant medication.

Prior to admission, James had resided in a foster home, his fifth placement in as many years. He became unusually unmanageable upon the arrival of a younger foster brother. James had last seen his biological mother three years prior to admission.

While James was considered to be highly impulsive and inattentive, the extent to which those symptoms were secondary to pervasive social and emotional deficits was unclear. Also, there was a need to evaluate treatment with stimulant medication because previous attempts had not been systematically monitored.

James was evaluated with the GDS while off, on, then off medication. (Unfortunately, a double-blind, placebo approach could not be implemented.) James' performance while free of medication fell solidly in the Abnormal Range for both tasks. He demonstrated a clear inability to sustain attention and delay. The practioner noted that, while the scores were generally typical of the ADD/Hyperactive youngster, the test behaviors were somewhat uncharacteristic. Unlike many ADD/Hyperactive children, James appeared cooperative, anxious, eager to please, and concerned about his performance.

After a week of treatment with methylphenidate, James was retested and his scores showed marked improvement. His performances on the Delay and Vigilance Tasks were at or near the Normal range, as he demonstrated improved self-control and attentiveness. While he still displayed considerable immaturity, James nonetheless appeared calmer and more focused.

Two weeks later, James was tested once more, this time free of stimulant medication. Even though he boasted that he was an "expert" at these tests, his performance fell back to baseline levels. He was again unable to exert self-control on the Delay Task, and had a high number of commissive errors on the Vigilance Task.

It was suggested that attention deficits were embedded in a constellation of other difficulties, and that attentional aspects of his problems were responsive to pharmacotherapy. The recommendation was made to continue stimulant medication therapy as part of a comprehensive treatment program.

CONCLUSION

The use of computerized assessment in the evaluation of children referred for ADD/Hyperactivity can make for a more accurate, efficient diagnostic process. It allows the practitioner to incorporate data based on the child's actual behavior. Clinicians who pursue this approach need to select a procedure that is practical, rugged, well-normed, and energetically supported by research. A legitimate concern is that these computer-generated tasks present a relatively narrow range of demands, stimuli, and response modalities. While the ADD child's impulsivity is viewed as generally pervasive across areas of functioning, no security exists that performance on select computerized tasks will necessarily reflect the overall level of impulsivity for a particular child. In all situations, these objective data must be integrated into a comprehensive evaluation of the sort syggested by Barkley (1981, 1985). As part of a sophisticated clinical protocol, a computerized attention task can significantly enhance the diagnostic process.

References

Atkinson, A.W., Cohen, P.C. & Kelly, P.C. (1985, April). Attention deficit disorder: The effects of ritalin on self-esteem a comparison of ACTeRS teacher scale, Conner's parent scale, and Gordon Diagnostic system in diagnosis and managment. Paper presented at the American Academy of Pediatrics Meeting, Atlant, Georgia.

Barkley, R.A. (1981). Hyperactive children: A handbook for diagnosis and treatment. New York: The Guildford Press.

Barkley, R.A. (1985). Development of a multi-method clinical protocol for assessing stimulant drug responding in ADD children. Manuscipt submitted for publication.

Conners, C.K. (1980). Continuous performance test [Computer program]. Kensington, Maryland: Behavioral Medicine, Inc.

Gordon, M. (1979). The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive children. Journal of Abnormal Psychology, 7, 317-326.

Gordon, M. (1982). The Gordon Diagnostic System. .Littleton, Colorado: Clinical Diagnostics, Inc.

Gordon, M. & McClure, F.D. (1983). [The assessment of ADD/Hyperactivity in a public school population]. Unpublished raw data.

Gordon, M. (1984,1985). ADD/Hyperactivity Newsletter. (Available from Clinical Diagnostics, Inc., Educational Services, 300 East Mineral Avenue, #6, Littleton, Colorado 80122).

Greenberg, L.M. (October,1985). An objective measure of response to methylphenidate: Clinical validation of the VIRTEST. Paper presented at the 32nd Annual Meeting of the American Academy of Child Psychiatry. San Antonio, Texas.

Klee, S.H. & Garfinkel, B.D. (1983). The computerized continuous performance task: A new measure of inattention. Journal of Abnormal Child Psychology, 11(4), 489-496.

McClure,F.D. & Gordon, M. (1984). The performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. Journal of Abnormal Child Psychology, 12(4), 561-572.

Rapport, M.D., Dupaul, G.J., Kelly, K. & Jones, J. (1985). Comparing classroom and clinic measures of ADD-H: Dose response effects. Manuscript submitted for publication.

Rosvold, H.E., Mirsky, A.F., Sarason, I., Bransome, E.D. & Beck, L.H. (1956). A continuous performance test of brain damage. Journal of Consulting Psychology, 20, 343-350.

DEPARTMENT OF HEALTH AND HUMAN SERVICES **Public Health Service**

Food and Drug Administration
Center for Devices and
Radiological Health
8757 Georgia Avenue
Silver Spring, MD 20910

APRIL 30, 1986

CLINICAL DIAGNOSTICS, INC.
ATTN: BRUCE MEYER
300 E. MINERAL AVENUE, #6
LITTLETON, CO 80122

Ref : K854903
Product : GORDON DIAGNOSTIC SYSTEM GDS DATA ANALYSIS PROGRAM

We are holding your above-referenced Premarket Notification (510(k)) for 30 days pending receipt of the additional information that was requested by the Office of Device Evaluation. This information should be submitted in duplicate to:

Food and Drug Administration
Center for Devices and
Radiological Health
Document Mail Center (HFZ-401)
8757 Georgia Avenue
Silver Spring, Maryland 20910

When your additional information is received by the Office of Device Evaluation the 90-day period will begin again.

If after 30 days the requested information is not received, we will stop reviewing your submission. Pursuant to 21 CFR 20.29, a copy of your 510(k) submission will remain in the Office of Device Evaluation. If you then wish to resubmit this 510(k) notification, a new number will be assigned and the 90-day time period will begin again.

If you have procedural or policy questions, please contact the Division of Small Manufacturers Assistance at their toll-free number (800) 638-2041 or me at (301) 427-8162

Sincerely yours,

Robert I. Chissler
Premarket Notification Coordinator
Office of Device Evaluation
Center for Devices and
Radiological Health

DO NOT REMOVE THIS ROUTE SLIP!!!!

K-85-4903 4/30/86

FROM: CLINICAL DIAGNOSTICS, INC. ATTN: BRUCE MEYER 300 E. MINERAL AVENUE, #6 LITTLETON, CO 80122	LETTER DATE 11/25/85 TYPE OF DOCUMENT: 510 (k)	LOGIN DATE 12/09/85	DUE DATE 05/21/86 CONTROL # K854903
TO: ODE/DMC	CONT. CONF.: ? STATUS : H REV PANEL : ~~HO~~ NE PAN/PROD CODE(S): ~~HO~~/NE / /		
SUBJECT: GORDON DIAGNOSTIC SYSTEM GDS DATA ANALYSIS PROGRAM			
DECISION: DECISION DATE: / /	RQST INFO DATE: 02/10/86 DATE: 04/30/86 DATE: / /	INFO DUE DATE: 03/12/86 DATE: 05/30/86 DATE: / /	

SUPPLEMENT: 01 LTR DATE: 860219 LOGIN DATE: 860220

DEPARTMENT OF HEALTH & HUMAN SERVICES — Public Health Service

Memorandum

Date 28 April 1986

From REVIEWER(S) - NAME(S) Mattan / Hinckley

Subject 510(k) NOTIFICATION K854903/A

To THE RECORD

It is my recommendation that the subject 510(k) Notification:

_____ (A) Is substantially equivalent to marketed devices.

_____ (B) Requires premarket approval. NOT substantially equivalent to marketed devices.

X (C) Requires more data.

_____ (D) Is an incomplete submission. (See Submission Sheet).

Additional Comments:
Hold for additional information.
See enclosed phone memo.

The submitter requests:

X No Confidentiality

_____ Confidentiality for 90 days

_____ Continued Confidentiality exceeding 90 days

Class Code w/Panel: _____

REVIEW: __________
(BRANCH CHIEF) (DATE)

FINAL REVIEW: __________
(DIVISION DIRECTOR) (DATE)

DEPARTMENT OF HEALTH & HUMAN SERVICES Public Health Service

Memorandum

Date MAR 10 1986

From REVIEWER(S) - NAME(S) Mattan

Subject 510(k) NOTIFICATION K854903/A

To THE RECORD

It is my recommendation that the subject 510(k) Notification:

✓ (A) Is substantially equivalent to marketed devices.

____ (B) Requires premarket approval. NOT substantially equivalent to marketed devices.

____ (C) Requires more data.

____ (D) Is an incomplete submission. (See Submission Sheet).

Additional Comments:

MEDICAL COMPUTERS AND SOFTWARE

See Memo

Use old letter

The submitter requests:

✓ No Confidentiality

____ Confidentiality for 90 days

____ Continued Confidentiality exceeding 90 days

Class Code w/Panel:

80 LNX unclassified

REVIEW: [signature] (BRANCH CHIEF) 3/12/86 (DATE)

FINAL REVIEW: [signature] (DIVISION DIRECTOR) 3/14/86 (DATE)

MEMORANDUM OF TELEPHONE CONVERSATION

Between: Mr. Bruce G. Meyer Jr.
President
Clinical Diagnostics, Inc.
300 E. Mineral Ave. #6
Littleton, Colorado 80122
(303) 795-0438

And: Physiologist, DANRD, HFZ-430

Date: 28 April 1986

Subject: Premarket Notification K854903A

Mr. Meyer and I discussed the information he submitted in the above noted submission concerning the Gordon Diagnostic System (GDS). During our conversation he said he would forward the following information to HFZ-401 labeled as an addendum to the file:

1.) A description of the following tests; Continuous Performance Test (CPT), Delayed Reaction Time Test (DRTT), and the Differential Reinforcement of Low Rates (DRL), ~~including~~ a description of the devices or mechanisms used to perform the tests.

2.) Data substantiating the use of the above tests for screening or testing for hyperactivity or Attention Deficit Disorders.

The above data and descriptions will be in the form of literature published in professional or psychology journals.

3.) Data describing the normal population or population used to generate the normative data that accompanies the GDS and that is used to rate/score a patient. This will include a description of the population, size, age range, geographical locales, statistics, and how the data is used to provide normative data.

4.) Data substantiating the usefulness of GDS in pharmacotherapy as well as a description of the studies used for gathering the data.

Mr. Meyer said that the Distractibility Test, which was briefly mentioned in one of the enclosed articles, is not to be considered in this submission. When data sufficient to support the usefulness of that test is gathered a separate submission will be made to FDA describing that test.

Stephen M. Hinckley

MEMO RECORD	AVOID ERRORS PUT IT IN WRITING	DATE MAR 12 1986
FROM: Biomedical Engineer		OFFICE ODE
TO: Record		DIVISION DGGD
SUBJECT: Clinical Diagnostics Inc. GDS and DAP		

SUMMARY

The Gordon Diagnostic System (GDS) is a portable electronic device designed to assess deficits in attention and impulse control in children. It has been developed for use by clinicians as an aid in the diagnosis of attention deficit disorders ~~with~~ as well as some forms of learning disabilities. The GDS administers two tasks: 1) The Delay Task measures a child's ability to refrain from pressing in order to win points and; 2) the Vigilance Task assesses how well a child sustains attention over a long period of time.

The GDS Data Analysis Program (DAP) is an optional ~~software~~ program designed to simplify the collection of data generated by the GDS, by organizing and storing the data.

A child is placed in one of three categories (normal, borderline, abnormal) based on the comparison of his scores to the norms established through a standardization study.

The two tasks performed in the GDS are based on psychological testing instruments developed for research. The Continuous Performance Test ('56) is the basis for the GDS Vigilance Task, and the Differential Reinforcement of Low Rate Respon-

SIGNATURE

DOCUMENT NO. K854903

FORM FDA 2034 (12/80)

ding ('67) is the basis for the GDS Delay Task.

Amalie C. Mattan

DEPARTMENT OF HEALTH AND HUMAN SERVICES **Public Health Service**

Food and Drug Administration
Center for Devices and
Radiological Health
8757 Georgia Avenue
Silver Spring, MD 20910

FEBRUARY 20, 1986

CLINICAL DIAGNOSTICS, INC.
ATTN: BRUCE MEYER
300 E. MINERAL AVENUE, #6
LITTLETON, CO 80122

D.C. Number : K854903
Received : 02-20-86
Product : GORDON DIAGNOSTIC SYSTEM GDS DATA ANALYSIS PROGRAM

The additional information you have submitted has been received.

-- We will notify you when the processing of your submission has been completed or if any additional information is required. You are required to wait ninety (90) days after the received date shown above or until receipt of a "substantially equivalent" letter before placing the product into commercial distribution. I suggest that you contact us if you have not been notified in writing at the end of this ninety (90) day period before you begin marketing you device. Written questions concerning the status of your submission should be sent to:

Food and Drug Administration
Center for Devices and
Radiological Health
Office of Device Evaluation
Document Mail Center (HFZ-401)
8757 Georgia Avenue
Silver Spring, Maryland 20910

If you have procedural or policy questions, please contact the Division of Small Manufacturers Assistance at their toll-free number (800) 638-2041 or me at (301) 427-8162

Sincerely yours,

Robert I. Chissler
Premarket Notification Coordinator
Office of Device Evaluation
Center for Devices and
Radiological Health

- NEW ADDRESS -

300 E. Mineral Avenue, #6
Littleton, Colorado 80122
303-795-0438

K854903/A

Clinical Diagnostics, Inc.

BASIS MEDICAL PRODUCTS

February 19, 1986

RECEIVED 1986 FEB 20 AM 11:42

Mr. Robert Gatling
General Hospital and Personal
Use Devices
Food and Drug Administration
Document Mail Center (HFZ-401)
8757 Georgia Avenue
Silver Spring, Maryland 29010

Dear Mr. Gatling:

Attached is our response to the requests for additional information by telephone from Amalie. Please expedite action on our 510K application as quickly as possible. Clinical Diagnostics has been marketing the Gordon Diagnostics System for the past two years, and based on directives from your department we are now holding all orders. I would greatly appreciate a telephone call as to the current status.

Thank you very much for your kind cooperation.

Sincerely,

CLINICAL DIAGNOSTICS, INC.

Bruce G. Meyer, Jr.
President

BGM:em

Clinical Diagnostics, Inc.

Ref: K854903
Product: Gordon Diagnostic System
GDS Data Anaylsis Program

In response to the request for additional information:

(b)(4)

[redacted]

1536 COLE BOULEVARD, SUITE 350, GOLDEN, CO 80401
(303) 232-7129
(800) 521-4503

(b)(4)

[redacted]

Sincerely,

CLINICAL DIAGNOSTICS, INC.

Bruce G. Meyer, Jr.
President

BGM:em

DEPARTMENT OF HEALTH AND HUMAN SERVICES **Public Health Service**

Food and Drug Administration
Center for Devices and
Radiological Health
8757 Georgia Avenue
Silver Spring, MD 20910

FEBRUARY 10, 1986

CLINICAL DIAGNOSTICS, INC.
ATTN: BRUCE MEYER
300 E. MINERAL AVENUE, #6
LITTLETON, CO 80122

Ref : K854903
Product : GORDON DIAGNOSTIC SYSTEM GDS DATA ANALYSIS PROGRAM

We are holding your above-referenced Premarket Notification (510(k)) for 30 days pending receipt of the additional information that was requested by the Office of Device Evaluation. This information should be submitted in duplicate to:

Food and Drug Administration
Center for Devices and
Radiological Health
Document Mail Center (HFZ-401)
8757 Georgia Avenue
Silver Spring, Maryland 20910

When your additional information is received by the Office of Device Evaluation the 90-day period will begin again.

If after 30 days the requested information is not received, we will stop reviewing your submission. Pursuant to 21 CFR 20.29, a copy of your 510(k) submission will remain in the Office of Device Evaluation. If you then wish to resubmit this 510(k) notification, a new number will be assigned and the 90-day time period will begin again.

If you have procedural or policy questions, please contact the Division of Small Manufacturers Assistance at their toll-free number (800) 638-2041 or me at (301) 427-8162

Sincerely yours,

Robert I. Chissler
Premarket Notification Coordinator
Office of Device Evaluation
Center for Devices and
Radiological Health

DO NOT REMOVE THIS ROUTE SLIP!!!!

K-85-4903 2/10/86

FROM: CLINICAL DIAGNOSTICS, INC. ATTN: BRUCE MEYER 300 E. MINERAL AVENUE, #6 LITTLETON, CO 80122	LETTER DATE 11/25/85	LOGIN DATE 12/09/85	DUE DATE 03/09/86
	TYPE OF DOCUMENT: 510 (k)		CONTROL # K854903
TO: ODE/DMC	CONT. CONF.: ? STATUS : H REV PANEL : HO PAN/PROD CODE(S): HO/ / /		
SUBJECT: GORDON DIAGNOSTIC SYSTEM GDS DATA ANALYSIS PROGRAM			
DECISION: DECISION DATE: / /	RQST INFO DATE: 02/10/86 DATE: / / DATE: / /	INFO DUE DATE: 03/12/86 DATE: / / DATE: / /	

DEPARTMENT OF HEALTH & HUMAN SERVICES Public Health Service

Memorandum

Date FEB 7 1986

From REVIEWER(S) - NAME(S) Mittan

Subject 510(k) NOTIFICATION K854903

To THE RECORD

It is my recommendation that the subject 510(k) Notification:

_____ (A) Is substantially equivalent to marketed devices.

_____ (B) Requires premarket approval. NOT substantially equivalent to marketed devices.

__✓__ (C) Requires more data.

_____ (D) Is an incomplete submission. (See Submission Sheet).

Additional Comments:

Hold

The submitter requests:

__✓__ No Confidentiality

_____ Confidentiality for 90 days

_____ Continued Confidentiality exceeding 90 days

unclassified
Class Code w/Panel: 80 LNX

REVIEW: ______________________________
(BRANCH CHIEF) (DATE)

FINAL REVIEW: ______________________________
(DIVISION DIRECTOR) (DATE)

MEMO RECORD	AVOID ERRORS PUT IT IN WRITING	DATE FEB 7 1986 10:40AM
FROM: Biomedical Engineer		OFFICE ODE
TO: Record		DIVISION DGGD
SUBJECT: K854903 Gordon Diagnostic System GDS DAP		

SUMMARY

I spoke with Mr. Bruce Meyer, Jr., President, to request the following additional information:

1. The similarities and differences between the GDS and the products to which Clinical Diagnostics is claiming S.E (Continuous Performance test, CPT; and the Differential Reinforcement of Low Rate Responding, DRL);
2. The basis for the educational recommendations; and
3. a description of and a basis for the "suggestions for pharmacotherapy" (Mr. Meyer said that this will not be a feature of the GDS DAP, and will include a statement to that effect ~~to~~ with the additional information)

Mr. Meyer will send the additional info. to DMC and is aware that the file will be on hold until that info is received.

SIGNATURE	DOCUMENT NO.
Amalie C. Matton	K854903

FORM FDA 2034 (12/80)

DEPARTMENT OF HEALTH & HUMAN SERVICES Public Health Service

Food and Drug Administration
8757 Georgia Avenue
Silver Spring MD 20910

DEC 12 1985

Clinical Diagnost.cs, Inc.
ATTN: Bruce G. Meyer, Jr.
300 E. Mineral Avenue, #6
Littleton, CO 80122

D.C. Number : K854903
Received : 12-9-85
Product : Gordon Diagnostic System
GDS Data Analysis Program

The Premarket Notification you have submitted as required under Section 510(k) of the Federal Food, Drug and Cosmetic Act for the above referenced device has been received and assigned a unique document control number (D.C. Number above). Please cite this D.C. Number in any future correspondence that relates to this submission.

We will notify you when the processing of this submission has been completed or if any additional information is required. You are required to wait ninety (90) days after the received date shown above or until receipt of a "substantially equivalent" letter before placing the product into commercial distribution. I suggest that you contact us if you have not been notified in writing at the end of this ninety (90) day period before you begin marketing your device. Written questions concerning the status of your submission should be sent to:

Food and Drug Administration
Center for Devices and
Radiological Health
Office of Device Evaluation
Document Mail Center (HFZ-401)
8757 Georgia Avenue
Silver Spring, Maryland 20910

If you have procedural or policy questions, please contact the Division of Small Manufacturers Assistance at their toll-free number (800) 638-2041 or me at (301) 427-8162.

Sincerely yours,

Robert I. Chissler
Premarket Notification Coordinator
Office of Device Evaluation
Center for Devices and
Radiological Health

K854903

- New Address -
300 E. Mineral Avenue, #6
Littleton, Colorado 80122
(303) 795-0438

Clinical Diagnostics, Inc.

November 25, 1985

Food and Drug Administration
Bureau of Medical Devices
Document Control Center (HFK-20)
8757 Georgia Avenue
Silver Spring, Maryland 20910

Re: 510K Notification

Attention: Document Control Clerk

This is to notify you of the intention of Clinical Diagnostics, Inc. to manufacture and market the following device:

Classification Name: Not Known

Common/Usual Name: Psychological Testing Device

Proprietary Name: Gordon Diagnostic System
GDS Data Analysis Program

Establishment Registration Number: 1721005

Classification: We are not aware that this device has been classified by FDA.

Performance Standard: None established under Section 514

Labeling/Promotional Material: Photographs and copies of promotional literature and manuals are enclosed.

Substantial Equivalence: We are not aware of any devices exactly like the Gordon Diagnostic System (GDS) and the GDS Data Analysis Program (GDSDAP). The two tasks performed in the GDS are based on psychological testing instruments developed for research. The Continuous Performance Test (CPT; Rusvold and Mirski, 1956) is the basis for the GDS Vigilance Task. The Differential Reinforcement of Low Rate Responding (DRL; Weisberg and Tragakis, 1967) is the basis for the GDS Delay Task.

The Gordon Diagnostic System Data Analysis Program is software to collate and analyze data produced by the GDS. It is not necessary for the operation of the GDS but serves as an accessory to facilitate data handling.

Respectfully submitted,

Bruce G. Meyer, Jr.
President

1536 COLE BOULEVARD, SUITE 350, GOLDEN, CO 80401 (303) 232-7129 (800) 521-4503

Clinical Diagnostics, Inc.

ABOUT THE GORDON DIAGNOSTIC SYSTEM

The Gordon Diagnostic System (GDS) is a tool to aid your diagnosis of children suspected of Attention Deficit Disorder (ADD). The instrument does this objectively by measuring two concerns of ADD: impulsivity and attentiveness. Game-like tasks, the Delay Task based on a DRL schedule, and the Vigilance Task based on a CPT schedule, result in recorded data that is compared to normed threshold tables.

DELAY TASK

The Delay Task is a self-paced task which measures the child's ability to surpress or delay impulsive behavioral responses. The task embodies a variety of processes, such as cognitive skills, behavioral demands, and motivational factors.

During the Delay Task, the child is instructed to press the button, wait a while, then press the button again. If the button is pressed too soon, no reward (point) will be obtained. Thus, the child must develop a method of determining the minimum time he must refrain from responding to obtain a point. He must develop a strategy for estimating each interresponse interval throughout the task.

The delay strategy employed involves a behavioral skill. Most nonimpulsive children use a "covert" strategy. That is, some form of internal, cognitively-oriented delaying procedure is employed. Usually, this is counting silently. Experience has shown that this method is quite effective, and children who use this method are quite adept at refraining from impulsive responses. Also, children who employ more "overt", or behaviorally manifested strategies, tend to perform much less efficiently on this task.

The Delay Task reflects the child's overall motivation to "do well". Since the task demands the application of cognitive and behavioral skills, the child must be reasonably motivated to maintain good performance throughout the task. The desire to gain control over a situation and the willingness to persevere with a rote task, are important factors in successful coping.

The microprocessor records time and responses. Standard parameters of a 6 second delay and four 120 second blocks are normed on approximately 1,000 subjects. Data recorded is transferred to a patient record form and compared to threshold tables. The Delay Task places maximum demands upon the ability to delay or refrain from emitting nonreinforced responses.

Clinical Diagnostics, Inc.

ATTENTION DEFICIT DISORDER/HYPERACTIVITY BACKGROUND INFORMATION

PREVALENCE:

- o 3-5% of school-aged population are affected. According to U.S. census figures, there are 30-40 million school-aged children.
- o 30-40% of all children referred for professional help due to behavior problems come with a presenting complaint associated with ADD/Hyperactivity. By far, ADD is the most "popular" behavior disorder of childhood.
- o 20% of adopted children are referred for symptoms of ADD/Hyperactivity.

DEMOGRAPHICS:

- o Ratio of boys to girls averages 6:1.
- o Highest incidence in lower socioeconomic status groups (SES) (highest in lower-middle class).
- o IQ's are generally in the average range.

TYPICAL CHARACTERISTICS:

- o Impulsivity.
- o Inattentiveness.
- o Restlessness.
- o Academic Underachievement.
- o Social Skill Deficits:
  - -Immaturity
  - -Egocentricity
  - -Poor regard for consequences of behavior
  - -Poor awareness of self and others
  - -Poor peer acceptance
- o Emotional Immaturity.
- o Exaggerated Emotional Reactions.
- o Low Frustration Tolerance.
- o Physical Findings:
  - -Motor Incoordination
  - -Enuresis (Bed Wetting) 30% of cases
  - -Encopresis (Soiling ones pants) 40% of cases
  - -Increased Incidence of Allergies
  - -Increased Incidence of Upper Respiratory Infection (URI)/cold and Otitis Media (middle ear infection)
  - -Increased Incidence of Minor Physical Anomalies
  - -Encephalogram (EEG) -- No Consistent Findings

OUTCOME:

- o Motor overactivity diminishes.
- o Chronic delay in problem areas.
- o Continued social difficulties.
- o Peer problems.
- o Increased association with children with conduct disorders.
- o Increased delinquency.
- o Poor school achievement.
- o Increased grade repetition.
- o Increased failure to finish school.
- o Decreased college attendance.

EMPLOYMENT DIFFICULTIES:

- o Not rated as abnormal by employer, but lower employer satisfaction.
- o Attains lower employment level.
- o Increased job and residence changes.
- o Emotional difficulties.
- o Increased depression.
- o Lower self esteem.
- o Lower frustration tolerance.
- o Increased hysteria in female hyperactives.

PREDICTORS OF OUTCOME:

- o Low IQ.
- o Low SES.
- o Presence of aggression/conduct disorder.
- o Degree of peer acceptance.
- o Degree of parent psychopathology.
- o Response to earlier treatment.

FAMILY CHARACTERISTICS:

- o 33-60% of parents with ADD child display similar symptoms.

Fathers:
- o Increased incidence of ADD/Hyperactivity symptoms.
- o Increased alcohol abuse.
- o Increased conduct problems.
- o Increased depression.
- o Increased desertion of family.
- o Increased difficulty maintaining consistent employment.

Mothers:
- o Increased depression.
- o Increased incidence of family discord/divorce (especially in lower SES).
- o First pregnancy at earlier age than common.
- o Increased alcohol abuse.
- o Increased stress in parenting situations.

Siblings:

- o Increased incidence of ADD/Hyperactivity symptoms.
- o Increased conduct problems.
- o Increased learning disabilities.

ASSESSMENT PROCEDURES:

Parents:

- o Careful Interview. (Review of symptoms; development history; social history; look for signs of parental depression and other effects of child's symptoms on parents).
- o Conners Parent Checklist or Achenbach Child Behavior Check List (CBCL).
- o Home Situations Questionnaire.
- o Locke-Wallace Marital Adjustment Survey.

Child:

- o Interview.
- o Gordon Diagnostic System (GDS).
- o Achievement Test.
- o Matching Familiar Figures Test (optional).
- o Review of school testing.

School:

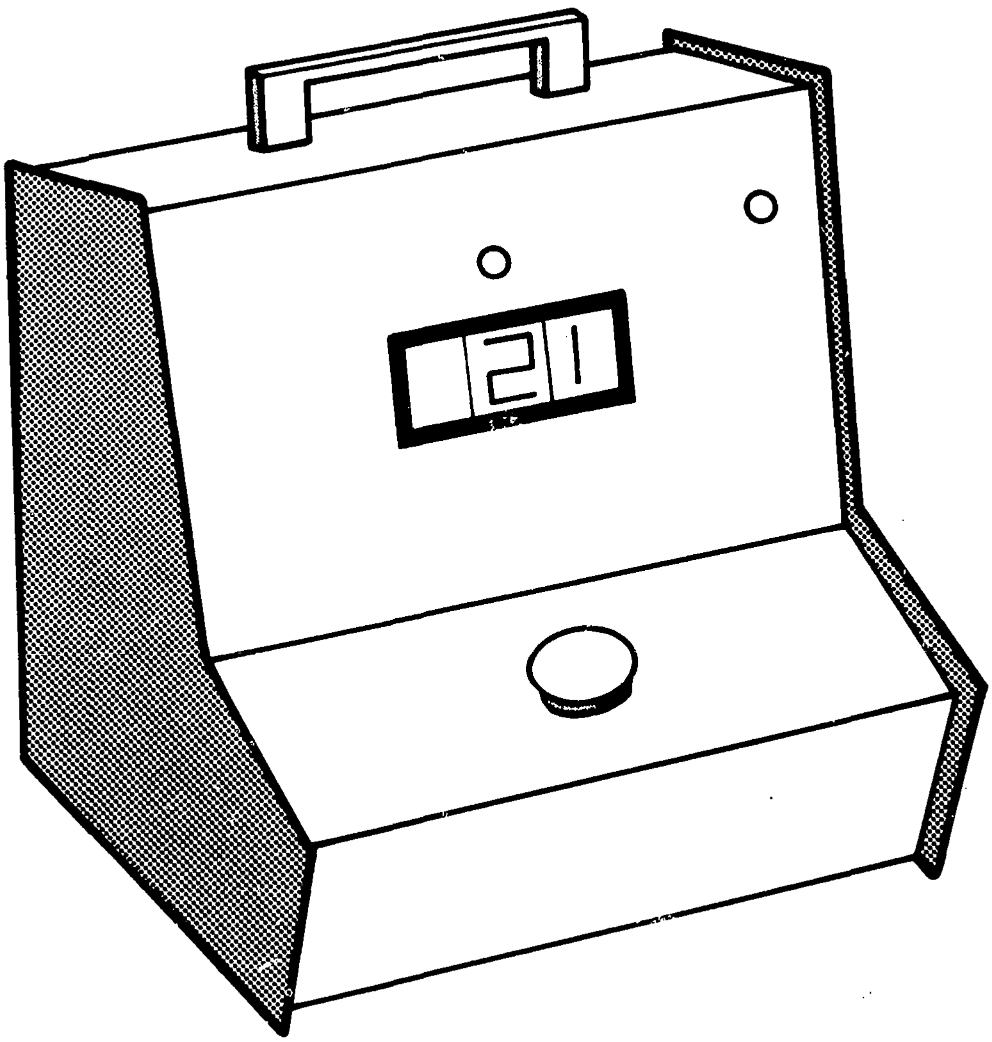
- o Discussion with teachers.
- o Conners Teacher Rating Form or Achenbach Teacher Report Form.
- o Kendall-Wilcox Self-Control Questionnaire.

MEDICATION TREATMENT:

- o Ritalin most commonly prescribed.
- o Typical dosage is 0.3mg - 0.5mg/kg.
- o Should not be prescribed if presence of tics.
- o Improves conduct at home and school.
- o Does not improve academic achievement (some controversy here).

THE GORDON DIAGNOSTIC SYSTEM

INSTRUCTION MANUAL



CLINICAL DIAGNOSTICS INC
1536 Cole Blvd., Golden, Colorado 80401
Suite 350 • 303-232-7129

Introduction

These instructions provide all necessary information to operate the GDS successfully. Although the device has been designed to facilitate proper administration, **read and follow the instructions carefully!!**

The Gordon Diagnostic System (GDS) is a portable electronic device designed to assess deficits in attention and impulse control in children. It has been developed for use by clinicians as an aid in the diagnosis of attention deficit disorders with or without hyperactivity, as well as some forms of learning disabilities. Research has shown that the GDS yields important information about a child's functioning and can differentiate accurately between youngsters classified as hyperactive and nonhyperactive.

The GDS administers two tasks:

(1) The Delay Task measures a child's ability to refrain from pressing in order to win points and;

(2) The Vigilance Task assesses how well a child sustains attention over a long period of time.

Both qualities -- the ability to delay and the ability to maintain attention -- are considered by authorities to be the main areas where children diagnosed hyperactive run into trouble. Such children find it very difficult to stop, look, listen and think before they try to solve a problem. They are impulsive and inattentive, just the traits which the GDS measures. (For further information refer to the GDS Interpretive Guide).

Getting Started...

Open the rear panel with the key provided. (Controls for the GDS are kept locked to prevent children from tampering with them during the operation of the tasks.) Plug the GDS into a 120 volt, 60 Hz grounded socket and turn on the main power switch, located in the rear panel. The front display should light up and you should also hear a tone. If the display does not light, check to make sure that the plug is firmly in the socket.

NOTE: If you want to turn the GDS quickly off then on again, you must wait a few seconds before pushing the switch back to the **"ON"** position. If you *do not wait long enough, the tone will not sound and* you will have to turn the machine off, wait three seconds, and turn it on again.

Delay Task

The Delay Task, which has traditionally been administered before the Vigilance Task, requires the child to wait a set period of time (called the Delay Interval) before pressing the blue response button. If the subject has waited long enough, the Reward Light will shine and the Reward Counter on the front panel will increment. If the child presses the button before the Delay Interval has elapsed then no point is earned, the light stays off and the Delay timer resets.

The GDS records Delay Task performance over four successive time Blocks. You can select the length of the Blocks (1-999 seconds) and the Delay Intervals (1-99 seconds) independently for each Block. These Blocks are internal recording units to monitor the child's performance along the session. The child is not shown when one Time Block stops and another begins.

Standardization data for the Delay Task were gathered using a 6-second Delay Interval and identical time Blocks of 120 seconds each (Standard Parameters). To program parameters into the GDS microprocessors:

1. Set the Task Selector Switch to **DELAY TASK.** (The position of the Mode Selector is irrelevant to this task).

2. Set the **INTERVAL** thumbwheel to the number of seconds you want the child to wait between each press of the button. The Standard Parameter is 6 seconds.

3. Set the **BLOCK LENGTH** thumbwheel to the desired number of seconds. The Standard Parameters call for **BLOCK LENGTHS** of 120 seconds each.

4. To load the parameters into the GDS microprocessor, press the **BLOCK 1** button of the Block **PROGRAMMER.** This communicates to the GDS the parameters you desire for Block 1. If you want to enter these same parameters for the other three blocks, press the Block **PROGRAMMER** buttons for Blocks 2, 3, and 4. Therefore, if you want the identical **DELAY INTERVAL** and **BLOCK LENGTH** for all four blocks, set the thumbwheels to the desired number of seconds and press the four Block **PROGRAMMER** buttons. A tone should sound each time a **PROGRAMMER** button is depressed.

5. If you want to set different parameters for each block, change the thumbwheels to the desired settings and press the Block Programmer buttons to load the parameters for a particular block.

NOTE: When the power to the GDS is turned on, it automatically reads and loads the parameters showing on the thumbwheels into all Blocks. If those are the desired settings, you can start at step #6.

NOTE: If you make an error in entering parameters, simply reset the thumbwheels and press the appropriate Block Programmer button(s).

6. **SET THE DATA SELECTOR TO POSITION #1. THE TASK WILL NOT OPERATE UNLESS THE DIAL IS IN THE FIRST POSITION.**

7. The GDS is now ready to begin the Delay Task. Subjects have traditionally been given the following instructions:

   You're going to play a game in which you will get a chance to win a lot of points. Do you see this light (pointing to the small red light)? Every time you make this light go on you'll earn a point and this counter (pointing) will keep track of how many points you've won. At the end of the game we'll see how many points you've earned. Now, to make the light go on all you have to do is push this blue button (pointing), and wait a little while before pressing it again. You just press this blue button, wait a while, then press it again. If you press it again too soon, though, you will not get a point, the light won't go on, and you'll have to wait a while before you can press it to get another point. But if you push the button, wait a while, then push it again, you'll get a point every time.

Que K854903 84 4 1

IMAGE EVALUATION TEST TARGET (MT-3)

150mm

6"

PHOTOGRAPHIC SCIENCES CORPORATION
770 BASKET ROAD
P.O. BOX 338
WEBSTER, NEW YORK 14580
(716) 265-1600

K854903

Delay Task (Cont.)

8. At this time you will want to answer any questions and make sure that the child understands the instructions.

9. The examiner then says:

   I'll sit here and wait until you're done. You'll know when the game is over because the green light will come on again. Please let me know when the green light comes on. If you have any questions or want to talk about the game, I want you to wait until after the game is over, and we can discuss it then.

10. When ready to begin the task, press the **RESET** button and then the green **START** button. You can do this either before locking the back flap or after by pushing the buttons with a pencil through the holes. You should hear the click of a button as each is depressed. If you hear the tone the task has begun and you should tell the child that the game has started. If you do not hear the tone either the **DATA SELECTOR** dial was not set to position #1 or you did not press the buttons hard enough.

11. Now that the task has begun, you should sit behind and somewhat to the side of the child so as not to be a distraction. Conversation should be discouraged. As with any testing situation, observations of the child's approach to the task can be highly informative.

12. When the task is over a tone sounds and the green Game Over light will come on. You may find it helpful to ask the child questions about his or her approach to the task. Subjects have typically been asked, "What did you do to help you wait?" and "Did you find it easy or hard to play the game?".

13. At this point you can put the unit aside and go on to the other tests. The data will be stored **as long as the unit is turned on** and plugged in. If you wish to administer the Vigilance Task, you must first record the Delay Task data.

14. **TO RECORD DATA** open back flap with the key. By turning the **DATA SELECTOR** to each position you will see on the **DATA DISPLAY** the total number of rewards (times the button was pressed after waiting long enough) and responses (times the button was pressed) for the whole session, as well as for each of the 4 Time Blocks. The key for the **DATA SELECTOR** positions is printed on the inside of the back flap and are as follows:

Data Selector Position	Data Displayed
1	Total Rewards
2	Total Responses
3	Block 1 Rewards
4	Block 1 Responses
5	Block 2 Rewards
6	Block 2 Responses
7	Block 3 Rewards
8	Block 3 Responses
9	Block 4 Rewards
10	Block 4 Responses

Make sure the Data Selector is set to position #1 when you are finished recording the data.

15. Once you have finished writing down the results you can either turn the unit off, move on to the Vigilance Task, or readminister the Delay Task. If you turn the power off all data is erased from the GDS microprocessor.

NOTE: If you wish to clear the front display, press the **RESET** button. If you have not turned off the power and want to readminister the Delay Task with the same parameters, simply press **RESET** and **START** to begin the task.

Delay Task

The Delay Task, which has traditionally been administered before the Vigilance Task, requires the child to wait a set period of time (called the Delay Interval) before pressing the blue response button. If the subject has waited long enough, the Reward Light will shine and the Reward Counter on the front panel will increment. If the child presses the button before the Delay Interval has elapsed then no point is earned, the light stays off and the Delay timer resets.

The GDS records Delay Task performance over four successive time Blocks. You can select the length of the Blocks (1-999 seconds) and the Delay Intervals (1-99 seconds) independently for each Block. These Blocks are internal recording units to monitor the child's performance along the session. The child is not shown when one Time Block stops and another begins.

Standardization data for the Delay Task were gathered using a 6-second Delay Interval and identical time Blocks of 120 seconds each (Standard Parameters). To program parameters into the GDS microprocessors:

1. Set the Task Selector Switch to **DELAY TASK.** (The position of the Mode Selector is irrelevant to this task).
2. Set the **INTERVAL** thumbwheel to the number of seconds you want the child to wait between each press of the button. The Standard Parameter is 6 seconds.
3. Set the **BLOCK LENGTH** thumbwheel to the desired number of seconds. The Standard Parameters call for **BLOCK LENGTHS** of 120 seconds each.
4. To load the parameters into the GDS microprocessor, press the **BLOCK 1** button of the Block **PROGRAMMER.** This communicates to the GDS the parameters you desire for Block 1. If you want to enter these same parameters for the other three blocks, press the Block **PROGRAMMER** buttons for Blocks 2, 3, and 4. Therefore, if you want the identical **DELAY INTERVAL** and **BLOCK LENGTH** for all four blocks, set the thumbwheels to the desired number of seconds and press the four Block **PROGRAMMER** buttons. A tone should sound each time a **PROGRAMMER** button is depressed.
5. If you want to set different parameters for each block, change the thumbwheels to the desired settings and press the Block Programmer buttons to load the parameters for a particular block.

NOTE: When the power to the GDS is turned on, it automatically reads and loads the parameters showing on the thumbwheels into all Blocks. If those are the desired settings, you can start at step #6.

NOTE: If you make an error in entering parameters, simply reset the thumbwheels and press the appropriate Block Programmer button(s).

6. **SET THE DATA SELECTOR TO POSITION #1. THE TASK WILL NOT OPERATE UNLESS THE DIAL IS IN THE FIRST POSITION.**
7. The GDS is now ready to begin the Delay Task. Subjects have traditionally been given the following instructions:

   You're going to play a game in which you will get a chance to win a lot of points. Do you see this light (pointing to the small red light)? Every time you make this light go on you'll earn a point and this counter (pointing) will keep track of how many points you've won. At the end of the game we'll see how many points you've earned. Now, to make the light go on all you have to do is push this blue button (pointing), and wait a little while before pressing it again. You just press this blue button, wait a while, then press it again. If you press it again too soon, though, you will not get a point, the light won't go on, and you'll have to wait a while before you can press it to get another point. But if you push the button, wait a while, then push it again, you'll get a point every time.

Vigilance Task

The Vigilance Task has two modes. In the "1/9" mode a series of numbers appears on the front display and the child is required to press the button every time a "9" appears immediately after a "1". The GDS records the number of Correct presses, the number of times the child did not press the button upon the appearance of a "1/9" combination (Omissions), and the number of improper presses (Commissions). The "1" task is similar, but for children under six years of age or for those who cannot follow the "1/9" instructions. The child is told to press the button every time a "1" appears.

The GDS records performance on the Vigilance Task over **three** Time Blocks (unlike four for the Delay Task). Normative data was derived using three identical Time Blocks of 180 seconds each.

The Presentation Interval is the time from when a number first appears to the initial presentation of the next digit. A Presentation Interval of one second was the Standard Parameter for the standardization studies. Numbers always appear for two tenths of a second (i.e. 200 milliseconds) regardless of the Presentation Interval setting. Consequently, when the Interval is set at 1 second, the digit appears for .2 seconds and the blanking time (when nothing appears on the screen) is .8 seconds. If you chose an **INTERVAL** of 2 seconds, the number appears for .2 seconds with a blanking time of 1.8 seconds.

Procedures for setting parameters and recording data for the Vigilance Task are very similar to those for the Delay Task:

1. Set the Task Selector Switch to **VIGILANCE TASK** and the Mode Selector either to "1/9" or "1". Clear the front display by pressing the **RESET** button.
2. Set the **INTERVAL** thumbwheel to the desired number of seconds you want between presentation of the digits. An **INTERVAL** of 1 second was used to establish normative data.
3. Set the **BLOCK LENGTH** thumbwheel to the desired number of seconds. The Standard Parameters call for Block Lengths of 180 seconds for each of the three time Blocks.
4. To enter the parameters into the GDS for Block 1, press the Block 1 button of the **BLOCK PROGRAMMER.** If you want to enter these same parameters for the other two Blocks, press the **BLOCK PROGRAMMER** buttons for Blocks 2 and 3 (there is no Block 4 in the Vigilance Task). Again, the procedure for establishing identical Presentation Interval and Block Length parameters for all three blocks is to turn the thumbwheels to the desired settings and press the first three **BLOCK PROGRAMMER** buttons.
5. If you want to set different parameters for each block, change the thumbwheels to the desired settings and press the Block Programmer buttons to register the parameters for a particular block.
6. **THE DATA SELECTOR MUST BE SET TO POSITION #1. THE TASK CANNOT RUN UNLESS THE DIAL IS IN THE FIRST POSITION.**
7. Close and lock the back flap. The GDS is now ready to administer the Vigilance Task. Children have typically been given the following instructions:

   Now we are going to play a different game. This time you are going to see numbers flash on the screen and I want you to press the blue button only when you see a "1" and then a "9". If a "9" shows by itself without a "1" coming right before it then do not press the button. Only press the button if you see a "9" coming right after a "1".

   Children performing on the "1" Mode are told simply to press the button every time the number "1" appears. With very young children it is important to first check that they can recognize the number "1".
8. Before proceeding you will want to answer any questions and make sure that the child understands the instructions.

Vigilance Task (Cont.)

9. The examiner then says:

   I'll sit back here and wait until you've finished. You will know when the game is over when the green light comes on. If you have any questions or want to talk about the game, I want you to wait until the game is over and we can discuss it then.

10. To begin the game press the **RESET** and **START** buttons making sure you hear each click as it is depressed. The task will begin as soon as the **START** button is pushed and the tone sounds. You will want to tell the child to begin immediately. If you do not hear the tone then check to make sure the Data Selector dial is in position # 1 and then firmly press the **START** button.

11. While the child is performing on this task you should again sit behind and to the side.

12. The task has been completed when the tone sounds and the green Game Over light comes on. The data for the Vigilance Task will be be stored as long as the power is turned on and the **RESET** button has not been pressed.

13. At this point you can put the unit aside and go on to the other tests. The data will be stored **as long as the unit is turned on** and plugged in. If you wish to administer the Delay Task, you must first record the Vigilance Task data.

14. **TO RECORD DATA** open back flap with the key. By turning the **DATA SELECTOR** to each position you will see on the **DATA DISPLAY** the total number of Correct responses for the whole session as well as the Correct responses, Omissions and Commissions for the 3 time Blocks. The key for the **DATA SELECTOR** positions is printed on the inside of the back flap and is as follows:

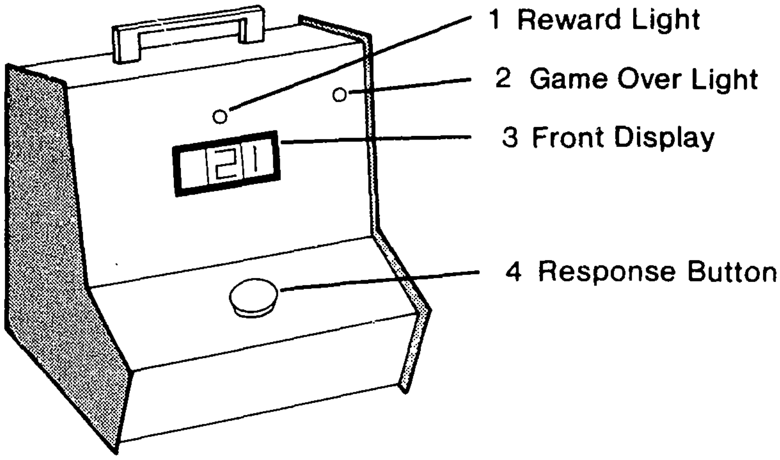
Data Selector Position	Data Displayed
1	Total Correct
2	Block 1 Correct
3	Block 1 Omissions
4	Block 1 Commissions
5	Block 2 Correct
6	Block 2 Omissions
7	Block 2 Commissions
8	Block 3 Correct
9	Block 3 Omissions
10	Block 3 Commissions

Make sure the **DATA SELECTOR** is set to position #1 when you are finished recording the data.

15. Once you have finished writing down the results you can either turn the unit off or repeat one of the tasks. If you turn the power off all data will be lost!

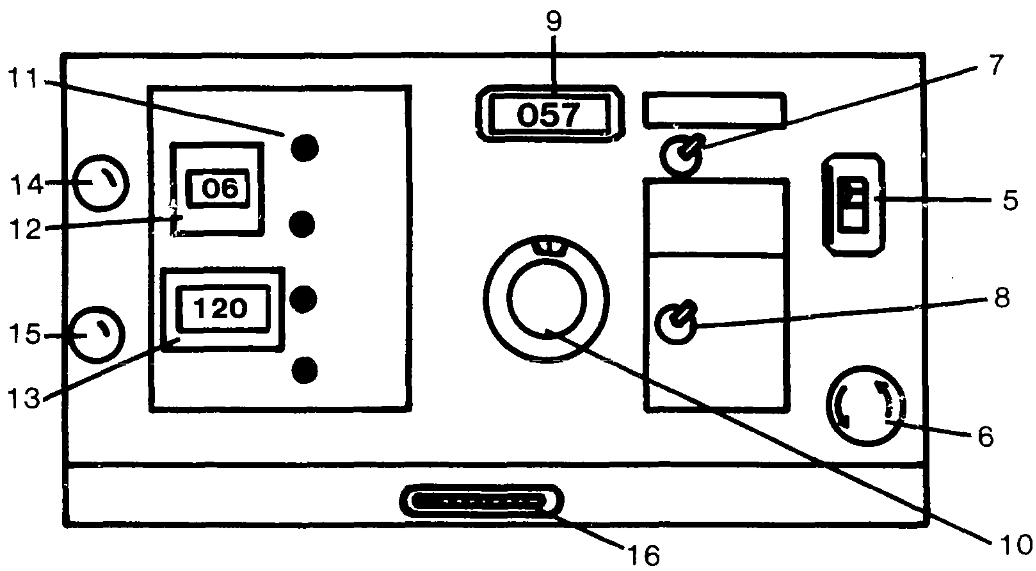
NOTE: If you wish to clear the front display, press the **RESET** button.

Description Of The GDS



Front View

1. Reward Light - Indicates rewarded response during Delay Task.
2. Game Over Light - Indicates end of Task.
3. Front Display - Counts rewards during Delay Task. Displays digit series during Vigilance Task.
4. Response Button - Pressed by child to perform Task.



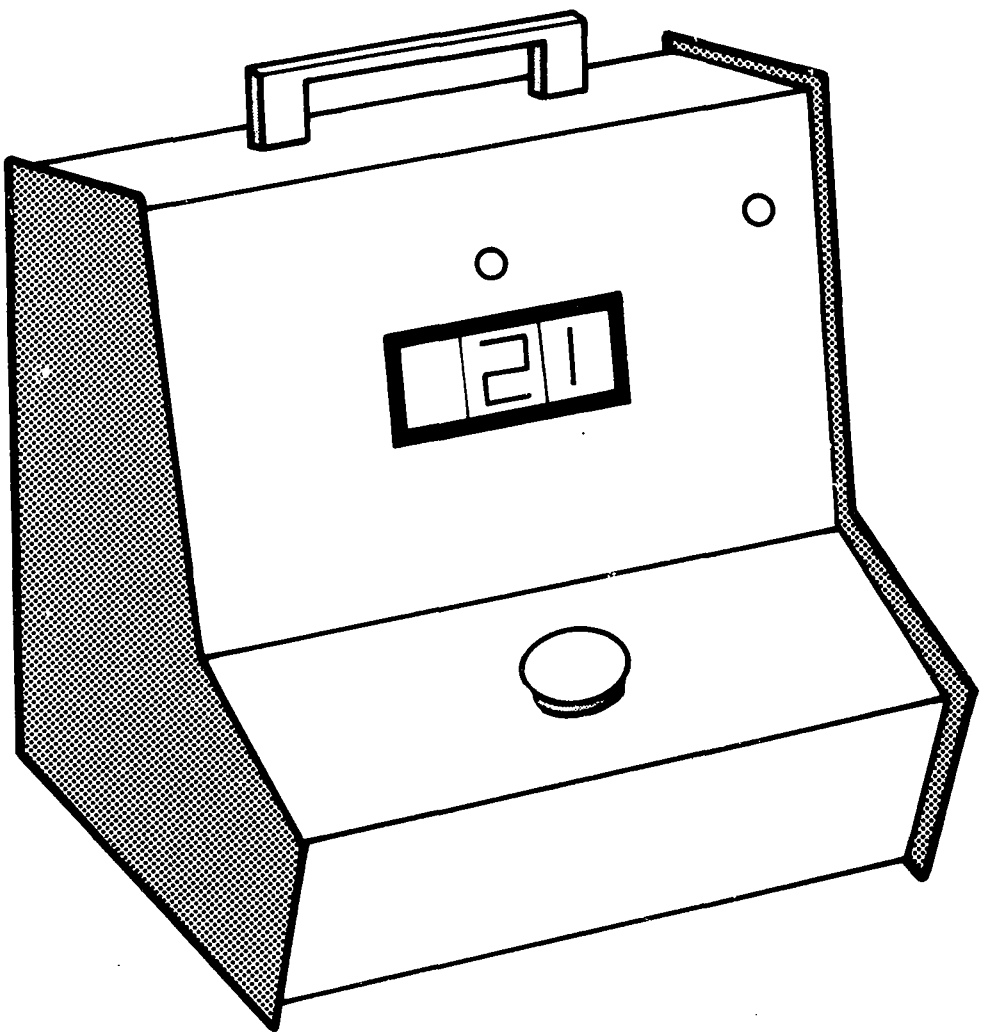
Rear View

5. Power Switch - Turns power on and off.
6. Fuse - Protects GDS from damage due to mechanical difficulties.
7. Task Selector Switch - Allows selection of Delay or Vigilance Task.
8. Mode Selector - Allows for selection between "1" and "1/9" version of Vigilance Task.
9. Data Display - Displays data from GDS performance.
10. Data Selector - Allows user to determine which data to display.
11. Block Programmer - Loads thumbwheel settings into GDS microprocessor.
12. Interval Thumbwheel - Sets Delay Interval for Delay Task and Presentation Interval for Vigilance Task.
13. Block Length Thumbwheel - Sets length of individual Time Block.
14. Reset Switch - Resets counters and clears displays.
15. Start Switch - Press to begin task.
16. RS-232 Communications Port - Enables GDS to link up with external computer terminals.

CLINICAL DIAGNOSTICS INC
1536 Cole Blvd., Golden, Colorado 80401
Suite 350 • 303-232-7129

THE GORDON DIAGNOSTIC SYSTEM

INTERPRETIVE GUIDE



CLINICAL DIAGNOSTICS INC
1536 Cole Blvd., Golden, Colorado 80401
Suite 350 • 303-232-7129

INTERPRETIVE GUIDE TO THE GORDON DIAGNOSTIC SYSTEM

Michael Gordon, Ph.D.
S.U.N.Y. Upstate Medical Center, Syracuse, New York

F. Daniel McClure, Ph.D.
D.C. Wilson Neuropsychiatric Hospital, Charlottesville, Virginia

Ernest M. Post, M.D.
S.U.N.Y. Upstate Medical Center, Syracuse, New York

Acknowledgments

The authors wish to express their gratitude to Ms. Jennifer DeLisi who, with great competence and cheer, coordinated our current research programs and assisted in the preparation of this Interpretive Guide. We also appreciate the fine efforts of Mrs. Mary Brown, Mrs. Noreen Lannon, and Miss Jennifer Kagan, and the crucial support of Mr. Kevin Bacher of the Liverpool School District, Liverpool, New York. These acknowledgments would not be complete without a note of gratitude to our families who have endured much throughout the development of the GDS. We are particularly thankful for the expertise and tolerance of Dr. Wendy Evers Gordon.

Aug. 17, 1983

Table Of Contents

INTRODUCTION TO THE INTERPRETIVE GUIDE

What is commonly referred to as hyperactivity probably raises more concerns than any other childhood behavior disorder of our time. It is also one of the most confusing and controversial issues ever to emerge in the arena of child mental health. The discussion of hyperactivity has highlighted the philosophical, theoretical, and professional biases among educators, pediatricians, psychiatrists, and psychologists.

How Common Is Hyperactivity?

Conservative estimates place the incidence of hyperactivity at between five and ten percent of the entire national elementary school population (Miller, Palkes, & Stewart, 1973). It is frequently at the top of the list when parents and educators are questioned about child behavior problems; it is by far the most frequent reason for consultations with child psychiatrists, pediatricians, and child psychologists; and it is a problem of major proportions in our nation's educational system. In a study of the Des Moines, Iowa public schools (Stone, Spence, Wilson, & Gibson, 1969) it was found that 53% of the boys and 30% of the girls were identified by their teachers as having symptoms of hyperactivity. Further, a national news report (Maynard, 1970) revealed that between five and ten percent of the 62,000 elementary school students in the Omaha, Nebraska public school system were being given behavior modifying drugs after being identified as hyperactive. These findings suggest a problem of epidemic proportions, but also reflect the confusion and misunderstanding surrounding the concept of hyperactivity. They dramatically speak to the paucity of valid, reliable diagnostic methods currently available in this area. In fact, accurate estimates of the true incidence of the disorder are not yet possible due to confusion over terminology and glaring inadequacies in diagnostic methods.

How Has Hyperactivity Traditionally Been Defined?

Much of this confusion arises from use of the term "hyperactivity". It has been very loosely applied to a wide variety of children who share a few common features; there has been little consensus among professionals as to exactly what the term implies. Some interpret it solely as a descriptive term referring to a child's overall level of motor output; others view it as reflecting the presence of neurological damage or "minimal brain dysfunction"; some link the term to certain autonomic dysfunctions such as arousal level or autonomic responsivity; still others consider it a disorder of higher cognitive structures, such as those involved in "metacognitive strategies" for problem solving. Since hyperactivity is not a specific disease, like diabetes, the term has been used in a variety of individually determined ways. Usually, a professional's view of hyperactivity is primarily determined by his or her own personal value system, theoretical orientation, professional training, and general philosophical view of development. It is no surprise, then, that highly subjective criteria have routinely been employed in "diagnosing" this disorder. As a result, the term has quickly lost any specificity which it might have otherwise embodied, and has essentially evolved into a rather meaningless "catch-all" category for many problem children. This situation has prompted some people to feel that hyperactivity simply represents "those aspects of a person's behavior which annoy the observer" (Buddenhagen & Sickler, 1969).

In response to this ambiguity, the DSM III psychiatric nomenclature has adopted the term "Attention Deficit Disorder With or Without Hyperactivity (ADD)". This change reflects the growing emphasis on evaluating the quality rather than the sheer quantity of the child's behavior. (The ADD classification will be used whenever possible in this manual. However, for the sake of continuity with the existing professional literature, the term hyperactivity is occasionally employed).

How Has The Disorder Been Diagnosed?

Richard Gardner (1979) noted that many professionals have routinely diagnosed the disorder solely on the basis of parent and/or teacher impressions, without actually having observed firm diagnostic signs themselves. This strongly contrasts with the usual practice of professionals, especially physicians, who rarely accept any other diagnosis made by a lay person. This sad state of affairs takes on added significance when one considers that psychostimulant medication has long been the treatment of choice for these youngsters and is still often used when the "diagnosis" is made. Physicians have frequently been the subject of heated indictments in professional journals, as well as in the lay press, for loose and indiscriminate prescription of these medications for supposedly affected children. Unfortunately, there have not been any effective alternatives to the subjective, highly biased, and often grossly inadequate methods of diagnosis. These inadequacies have often lead to ill-conceived and frequently inappropriate forms of treatment.

Under these circumstances, the need for a valid and objective diagnostic methodology cannot be overstated. Accurate diagnosis and assessment are as essential to effectively treating those who need it as they are to avoiding unnecessary treatment. If the basic assessment is erroneous or inadequate, then the effectiveness of treatment is seriously compromised.

Assessment Of ADD/Hyperactivity: A Conceptual Approach

Our approach to the evaluation of behaviors classified as hyperactive or ADD rests upon the functional assessment of the critical dimensions of impulse inhibition and attention. Our experience and a large body of research literature dictate that the essential features of the disorder lie in the inability to suppress behavioral responding and in the inability to focus and sustain attention. These processes are obviously very much interrelated. For example, it is difficult to expect a child to inhibit impulses to "act" when he or she cannot maintain attention for any length of time. Deficits in the ability to focus and sustain attention are therefore one of many causes for responses which are impulsive and poorly organized. By the same token, it is difficult to expect a child who is highly impulsive for other reasons and who cannot delay his or her responding to be able to attend to a task for any length of time. Their behavioral impulsivity seriously interferes with their ability to attend. This constellation of attentional deficits and lack of impulse inhibition have been referred to as the "stop, look, and listen" dimensions (Douglas, 1972), and they are the focus of our approach to the disorder. To reiterate, the two fundamental features involved in ADD are:

(a) the inability to refrain from emitting impulsive, and poorly directed behaviors, and
(b) the inability to focus attention and/or, sustain it over time.

Deficits in these two areas produce a deeply-ingrained pattern of behavior which we call the "impulsive style". This impulsive style is at the heart of what brings the child into conflict with his environment. We consider the term "hyperactivity" actually to refer to the symptoms of an impulsive style. The behavior of impulsive children may vary greatly from child to child, depending on a myriad of internal and external factors. Some children's impulsivity may manifest itself primarily at the level of attention and concentration and not in actual physical behavior. This child would most likely be diagnosed as having "ADD Without Hyperactivity" and require intervention aimed at developing stategies for sustaining attention and vigilance. The impulsive style of other children may also express itself through physical routes, and these children are often diagnosed as having "ADD with Hyperactivity". They require additional strategies for adopting less physical means of mediation and often are prescribed stimulant medication.

Emphasis on the impulsive style leads us away from focusing on specific behaviors or clusters of behaviors in evaluating these children. Instead, we have come to look for accurate assessment of the underlying processes of attention and delay which are at the core of the disorder. We assume that children with these deficits, i.e., who have an impulsive style, are at a significant disadvantage in adequately coping with problems that confront them.

It should be clear that our approach to the assessment of this disorder is not bound to any particular theoretical framework or conviction about etiology. Instead, we seek to describe, in an objective and realistic manner, the child's functional abilities and disabilities with the expectation that careful evaluation will lead to a better understanding of these problems and more efficient treatment planning. Indeed, the Gordon Diagnostic System (GDS) was developed with these goals in mind. We believe that it represents an accurate, stable, and objective approach to the assessment of the impulsive style and related disturbances of attention.

This interpretive guide provides further information on the background of the GDS and offers suggestions for interpreting the results of testings obtained with the device. Data from our research studies and clinical experience are presented, as are standardization tables for your use. To avoid gender prejudice, we have alternated male and female references throughout the manual.

THE ASSESSMENT OF ATTENTION DEFICIT DISORDERS

Background

Clinicians, educators, and physicians have been in chronic need of reliable and valid techniques for the assessment of attention deficit disorders (ADD) and impulsivity in children. The instruments that are currently available for the diagnosis of ADD fall into three basic categories: those actually intended for the measurement of other aspects of behavior which have been adapted to assess impulsivity, those which are largely limited to laboratory investigations, and questionnaires.

Most prominent among the instruments originally designed for the clinical assessment of other behavioral domains is the Wechsler Intelligence Scale for Children-Revised (WISC-R). While selected subtests or groups of subtests from the WISC-R are widely used in the identification of hyperactivity, studies have shown that they do not discriminate accurately between groups of hyperactive and nonhyperactive children (Douglas, 1972). Moreover, the WISC-R assesses a wide range of abilities and traits that are largely unrelated to impulsivity (Douglas, 1972). Thus, while all hyperactive children may perform poorly on these measures, there are many nonhyperactive children (e.g. learning disabled, or those with visual-motor deficits) who may also have difficulties with them for reasons other than impulsivity.

Among those instruments best suited for laboratory investigations, the most effective in differentiating hyperactive from nonhyperactive children are the Continuous Performance Test (Rosvold, Mirsky, Sarason, Bransone, & Beck, 1956), Delayed Reaction Time Test (Cohen, Douglas & Morganstern, 1972), and the Matching Familiar Figures Test (Kagan, 1966). While the Continuous Performance Test and the Delayed Reaction Time Test have been demonstrated to be excellent diagnostic measures, their technological sophistication and prohibitive expense make them unsuitable for widespread clinical use. The MFFT is quick, easy to administer, and has been demonstrated to be diagnostically sound. However, its reliability and validity have been the subject of considerable controversy which persists despite revisions of the measure (Egeland & Weinberg, 1976).

There are a variety of other measures available such as the Gardner Steadiness Tester (Gardner, 1979), Human Figure Drawings (Machover, 1949), certain configurations of Rorschach indicies, and paper-and-pencil "vigilance" tasks. As a group, these tend to suffer drawbacks stemming from impracticality of administration, limited information yield, and often poor standards of validity and reliability.

Behavioral inventories or questionnaires have also enjoyed widespread use, and many, such as the Teacher Rating Scale (Conners, 1969), the Child Behavior Check List (Achenbach, 1978), and the Self-Control Rating Scale (Kendall & Wilcox, 1979), have demonstrated good discriminative validity. However, there are serious inherent drawbacks to such a format, including rater bias, the possibility of "halo" effects, item overlap, vague rating criteria, issues of basic validity, and the fact that questionnaires yield no direct behavioral data.

There is thus an obvious and compelling need for a more straightforward, clinically applicable, and precise measure of hyperactive child's impulsivity. The Gordon Diagnostic System (GDS) shows considerable promise as a device which could help fill this gap in delivery of clinical service. The GDS is an easily-administered, game-like task which seems to precisely measure levels of impulsivity without contamination by other factors such as intelligence or visual-motor skills. It provides objective data on a particular child's ability to inhibit behavioral responding. Through observation of the child's performance and from responses to several post-test inquiries, important insights can be gained into a child's problem-solving style. This information can have strong bearing on treatment recommendations and educational planning. Thus, the GDS is unique because it yields not only quantitative data, but also a wealth of pertinent behavioral information which is crucial to the diagnostic process.

History Of The GDS

The Gordon Diagnostic System is an outgrowth of research by the senior author, Dr. Michael Gordon. While serving as a research assistant in the physiological psychology laboratory at The Ohio State University, Dr. Gordon was involved in a study of rats with lesions to the septal area of the brain (Gittis & Gordon, 1977). Following the ablations, these rats and a matched control of non-operated rats were tested on a Differential Reinforcement of Low Rates (DRL) schedule. The septal animals did poorly on this schedule because they lost the ability to adequately inhibit responding. At this same time, the literature on hyperactivity was pointing to impulsivity and attentional disorders as the core symptoms of hyperactivity (Douglas, 1972). It was also claimed that hyperactive children suffered from damage to the septal area. It therefore seemed sensible to investigate the performance of children classified as hyperactive on a customized version of the DRL paradigm (Gordon, 1979). Because of the promising results, the technique was further refined so that the procedure could become available to practicing clinicians. Most importantly, considerable reduction in the size of the equipment was made possible through the use of integrated circuits.

Description Of The GDS

The GDS is a portable, electronic unit which contains two tasks. The Delay Task is based upon a Differential Reinforcement of Low Rates (DRL) behavior schedule and requires the child to inhibit responding in order to gain a reward. Specifically, the child is instructed to press a button, wait, and then press the button again. If he refrains from responding for at least six seconds, a light flashes and a counter increments. If the child responds before the interval elapses then the timer resets and no reward points are recorded. The GDS Delay Task generates three major scores: the number of responses, the number of correct responses (i.e. Rewards) and the Efficiency Ratio which represents the percentage of correct responses.

The GDS also contains the Vigilance Task which is a version of the Continous Performance Test (Rosvold, et al., 1956). The child is presented with a series of digits on the front display and is told to press the button every time a "9" appears that had been preceded by a "1". The GDS records the number of correct presses, the number of times the child fails to press the button when the "1/9" combination appears (i.e. errors of omission), and the number of extraneous button presses (i.e. errors of commission). For the testing of children under 6 years of age, the GDS contains a "1" mode which requires the subject to push the response button every time a "1" appears. The same performance measures are recorded.

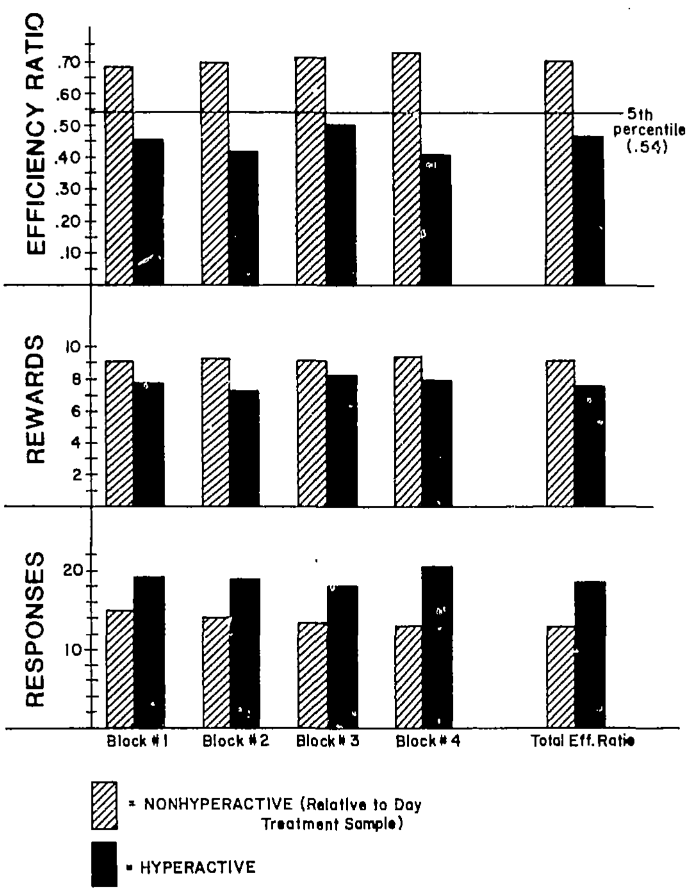
We have recently developed a GDS Data Analysis Program which enables a microcomputer to receive data directly from the GDS. The information can then be tabulated, graphed, and stored. The Data Analysis Program also allows measures of performance characteristics not available on the device itself (see "Ongoing Research").

Research Studies With The GDS

Data concerning the GDS Delay Task come from six studies conducted over as many years. The original experiment (Gordon, 1979) compared performances of 20 hyperactive and 20 nonhyperactive children selected from the patient roster of a child guidance clinic. There were highly significant differences between the two groups for all three GDS variables, i.e. Rewards, Responses and the Efficiency Ratio (Rewards/Responses). The children classified as hyperactive had profound difficulties in dealing successfully with the delay inherent in the task.

Since this initial work we have conducted several previously unreported studies on GDS performance. The original finding was recently replicated (McClure & Gordon, in press) with patients who had been placed in a day treatment program for seriously disturbed youngsters. Differences in GDS Delay Task scores between the 20 children classified as hyperactive and the 20 nonhyperactives in the comparison group were strikingly significant ($p < .0001$, see Figure).

DELAY TASK PERFORMANCE OF HYPERACTIVE VS. NONHYPERACTIVE SUBJECTS



To further judge the discriminative ability of the GDS, analyses were performed with groups formed on the basis of being identified as hyperactive or nonhyperactive by four criterion measures. These included the Teacher Rating Scale (Conners, 1969), the Self-Control Rating Scale (Kendall & Wilcox, 1979), the Teacher Report Form (Achenbach, 1980), and the Matching Familiar Figures Test - 20 (Cairns & Cammock, 1978). For subjects who were clearly identified as hyperactive (n = 8) or nonhyperactive (n = 8) on the basis of all four criteria, there was agreement between the GDS and the criterion measures in 15 cases (94%, Fisher's Exact Test, $p = .01$). When children were classified as hyperactive or nonhyperactive by **at least** three of the four criterion measures, there was agreement in 29 of 32 cases (91%, Fisher's Exact Test, $p = .001$).

Two additional studies appear to establish the validity of the GDS as a measure of impulsive/hyperactive behavior. In a group of 31 nonhyperactive subjects, GDS performance was significantly correlated with 12 teacher ratings. It was also related to impulsivity as measured by the Bender-Gestalt Test and the Rorschach Inkblots. Analysis of data from the files of 32 outpatient children also indicate strong relationships between GDS performance and parent ratings of impulsive behavior.

To establish norms for the GDS Delay Task scores, a standardization study was recently completed in the public schools. The sample included a total of 192 boys and girls in three age groups (6-7 yrs.; 8-9 yrs.; 10-11 yrs.). It was determined that the mean Efficiency Ratio for the Delay Task (i.e. Total Rewards/Total Response) was .84 ± .13 and that the scores were unaffected by age, sex, IQ and socioeconomic status. As with previous studies, performance remained consistent throughout the 8-minute session. In order to establish test-retest reliability, half the group was retested after 30-45 days, and it was found that there was a small learning effect that varied little among children.

Also completed recently was a study of 60 children who were classified by the school district as falling into one of four groups: attention deficit disorder, reading disabled, overanxious and normal. The GDS Delay and Vigilance Task successfully differentiated between the hyperactive group and the other three categories. It was also found that the correlations between the Delay Task Efficiency Ratio and the Vigilance Task scores fell in the .5 to .6 range of Pearson's **r**.

The Vigilance Task was first included in the GDS half-way through the standardization study. As such, there are now available norms for Vigilance Task performance. It was also used in the study of classified children and that data is currently being analyzed.

MEANS OF GDS VARIABLES: STANDARDIZATION SAMPLE

	6-7 year olds				8-9 year olds				10-11 year olds			
	Boys		Girls		Boys		Girls		Boys		Girls	
Delay Task (n=192)	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Total Rewards	40.0	9.0	36.6	12.0	42.8	10.5	42.2	10.4	47.3	10.5	45.9	8.4
Total Responses	52.6	15.2	47.1	17.7	53.2	14.9	52.0	14.2	54.4	13.7	53.8	9.8
Total Efficiency Ratio	.80	.16	.83	.18	.83	.13	.83	.13	.88	.09	.86	.11
Vigilance Task (n=91)	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Total Correct	30.7	9.4	25.7	11.8	37.6	6.0	38.5	3.7	41.8	1.7	41.5	2.3
Total Omissions	14.3	9.4	19.3	11.8	7.4	6.0	6.5	3.7	3.2	1.7	3.5	2.3
Total Commissions	16.1	15.5	9.3	5.0	11.8	10.7	9.6	10.3	6.9	8.0	5.1	4.1

Summary

While the GDS has successfully differentiated between diagnostic groups, the results are often difficult if not impossible to interpret because of the absence of reliable, independent criteria for making the classification of hyperactivity. Even when multiple criteria are employed the degree of diagnostic agreement among the criterion measures tends to be poor. Consequently, one can never be at all confident that misclassifications by the GDS according to other criteria are really errors of classification at all or are rather representative of more accurate assessment. Our experience to date has led us to gear future research away from criterion-based studies to those which focus on behavioral correlates of GDS performance.

CLINICAL APPLICATION OF THE GDS

INTRODUCTION

Effective utilization of the GDS rests upon the practitioner's understanding of the various processes involved when a child performs the tasks. The clinician must take into account the demands that the tasks place upon a child, the significance of the scores generated, and the value of the qualitative observations afforded during the performance of the tasks. The actual scores provide a basis for objectively comparing a child to a large group of non-ADD children. However, they must be interpreted with an understanding of the attributes and behaviors that affect a child's performance. Integrated with other observations, they can lead to making a diagnosis or classification and guide the pracitioner towards prescribing an environment that is appropriately designed to address the individual child's deficits and strengths.

In the following analysis, we will be dealing with abstract concepts such as "arousal", "attention", "mediation", and cognitive "processes". These are difficult to define and even more difficult to study objectively. We are presenting a large body of objective data derived from our experience testing a wide range of children with the GDS. However, much of what follows is the conceptual construct we have developed out of this experience.

THE GDS DELAY TASK

Demands Of The Task

The DelayTask has been demonstrated to measure a child's ability to suppress or delay impulsive behavioral responses. In spite of its apparent simplicity (see Description of the GDS on page 3), this task embodies a variety of processes which are quite complex. For this discussion, we have divided them into three groups:

1. Cognitive Skills
2. Behavioral Demands
3. Motivational Factors

Cognitive Skills

One of the major benefits of the GDS Delay Task is that a child's performance is relatively independent of measured IQ (See page 5). Nevertheless, a host of perceptual and cognitive skills are exercised in the performance of this task and need to be considered in the interpretation of the results.

Consider the situation which confronts the subject. He is instructed that, to win a point, he must "push the blue button, wait a while, then press it again", but that if the button is pressed prematurely, no reward will be obtained. These instructions are simple enough to be understood by most six-year olds and even some younger children. The reward is the illumination of a light and an incrementing number, both easily observed by a child of any age, even one with a moderate visual impairment.

As the duration of the interval is not mentioned in the intructions, the child must develop a method of determining the minimum time he must refrain from responding to obtain a point. Many non-impulsive children will find this interval through a series of successive approximations. That is, they will wait a relatively long period, press the button, be rewarded for waiting long enough, and then progressively shorten the duration of their delay time until they arrive at the approximate minimum interresponse time which will usually result in a reward. Some non-impulsive children will take the opposite strategy and gradually increase the time between responses until they find a successful interval. Other children will use a combination of both approaches to maximize their reward.

While this process of "fine tuning" responses seems simple, it actually constitutes a very sophisticated approach. The child must be able to efficiently utilize the feedback provided to guide his responses. He must also develop a strategy for estimating each interresponse interval throughout the remainder of the task.

Behavioral Demands

In addition to the cognitive ability required, successful performance of the task requires certain behavioral skills. While the child is determining the necessary interresponse interval, he must refrain from responding until the intended interval has elapsed. Since the reward is only obtained when the button is pressed, the self-restraint can be difficult. Once the initial strategy is devised, the child must maintain the suppression or delay of his responses for the appropriate interval throughout the remainder of the task.

The strategy used to repeatedly wait the appropriate interval involves a behavioral skill. Most non-impulsive children use a "covert" strategy, some form of internal, cognitively-oriented delaying procedure, usually counting silently. Our research has indicated that this is the most efficient method, in that children who employ this strategy are usually quite adept at refraining from impulsive responses. While the notion of counting to one's self may seem simple to an adult, it is actually a very sophisticated cognitive strategy for a child to devise independently. Our experience has shown that children who employ more "overt" or behaviorally manifested strategies tend to perform much less efficiently on this task.

Motivational Factors

While children are almost universally captivated by the task, the apparatus itself is designed to be simple, non-distracting and unrewarding. The motivation to perform well reflects both the child's overall motivation to "do well" and the child's receptivity to the task's rewards. Since the task demands the application of cognitive and behavioral skills, the child must be reasonably motivated to maintain a good performance throughout the task.

The subject is initially told that the task is a game in which he will have the chance to win a lot of points. This helps to instill in the subject a motivation to meet the challenge of "beating" the machine. A tone sounds to begin the game and sounds again at the end of the game along with the illumination of a green light. The brief illumination of the red light and the incremental counter function as the reward or "feedback" to provide the subject with reinforcement for a correct response.

The desire to gain control over a situation and the willingness to persevere with a rote task are important factors in successful coping. In fact, a respected theory of hyperactivity construes the problem as a developmental disorder of motivation (Glow & Glow, 1979). The notion is that these children are relatively unresponsive to feedback from the environment and come to see events in the world as generally unconnected to their own efforts. As a result, they are less aware of cause and effect relationships and fail to develop a sense of competence and intrinsic motivation. The motivation to succeed and the desire to master the situation are very key components of performance onthe Delay Task. Without the motivation to bring resources to bear, adequate coping is impossible.

In summary, there are a variety of processes which enter into efficiently performing the Delay Task. It requires the cognitive ability to devise and execute a strategy for waiting long enough to get rewarded. The child must have the behavioral control to suppress or delay the impulses to press the button too early. To achieve both of these, a substantial degree of motivation is needed. All of these processes must function in concert for the child to perform well. It is important to note that, even though the Delay Task is relatively long and mundane, it is not a measure of the ability to sustain attention, **per se.** Constant attention is not required and the child's interest is repeatedly renewed by the reward. In fact, the demands that the Delay Task places upon the attentional processes are minimal. A child could daydream, play, or be otherwise distracted during the Delay Task and still perform adequately on the same measure. He might even use those activities as strategies to achieve the necessary interresponse interval. Focusing and sustaining attention usually facilitate Delay Task performance, but are not essential. Rather, the Delay Task places maximum demands upon the ability to delay or to refrain from emitting non-reinforced responses.

Data From The GDS Delay Task

The GDS Delay Task generates a plethora of both quantitative and qualitative data about a child's performance. Although they have been described in previous sections, we have outlined below the major categories of information yielded by the task:

1. Total scores
   a. Rewards--Number of times response button pressed after Interval has elapsed (correct "hits")
   b. Responses--Total number of times button was pressed
   c. Efficiency Ratio = Rewards/Responses
2. Block Data - Rewards, Responses and Efficiency Ratio for each of 4 blocks
3. Task Monitoring Data (available using the GDS Data Analysis Program)
   Distribution of interresponse intervals
4. Subjective Data
   a. Strategy - "cognitive" vs. "physical"
   b. Bursts of responses
   c. Other unique responses

Interpretation Of The Delay Task

Interpreting the results of the delay task requires the practitioner to evaluate both the objective and subjective data. Evaluation of the objective data is facilitated by comparing results to the normative data (see page 5) and to results we have obtained testing subjects in clinical settings (see page 4). We have found that in normal children, the performance does not differ significantly among the blocks; there is no significant learning effect or deterioration with time. We therefore feel that you should first evaluate the total scores and only need to examine the block scores closely if the total scores suggest abnormal responses.

Evaluation of the subjective data is less definitive, but in the discussion that follows, we have presented an overview of our experience testing a variety of children. While the subjective information is invaluable for assessing the child and planning appropriate learning environments, your interpretation will be strongly affected by your professional training.

Evaluating Total Scores

In the allotted eight minutes of the delay task (Standard Parameters), there is the theoretical opportunity for 80 Rewards. Since there is an initial adjustment period and no one can time their interresponse interval precisely, a score of 80 points is unrealistic. Our experience has shown that most non-impulsive children obtain between 30 and 50 points out of 35 to 65 total responses, giving an efficiency ratio of .70 to .98.

The Efficiency Ratio has been demonstrated to be the best

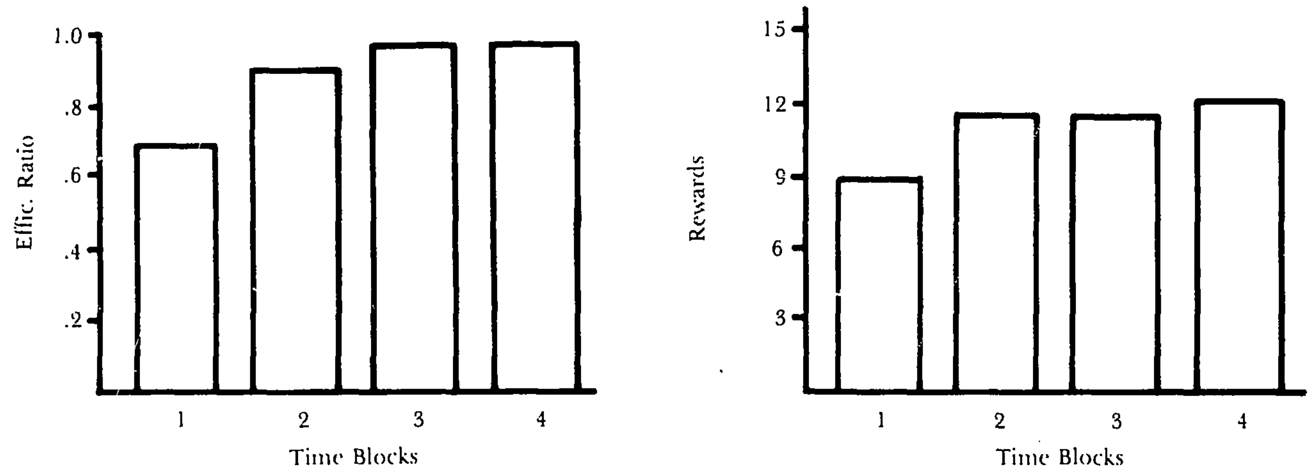
Delay Task index (see page 4) for discriminating between normal children and those judged to be hyperactive by a variety of other criteria. A child who cannot inhibit the frequent impulses to press the button will have a high number of total Responses. He will also tend to gain fewer total Rewards, because most of his responses will have occurred too soon after a previous impulsive response to be rewarded. As a ratio of Rewards to Responses, the Efficiency Ratio best reflects the effects of impulsivity on task performance.

However, the Efficiency Ratio can be misleading and needs to be interpreted in conjunction with the absolute number of Rewards. A child who presses the button only once every two minutes will have an Efficiency Ratio of 1.0, but is clearly different from the normal children we have tested. If a child has a normal Efficiency Ratio (> 0.55), but too few Rewards (< 18) several possibilites are evident. Most frequently, these scores are interpreted in terms of motivational factors such as resistance, acute stress or rebelliousness. Sometimes a child may recognize that he lacks the patterns of control necessary to perform well and overcompensate by ignoring the task for long periods.

Evaluating Block Scores

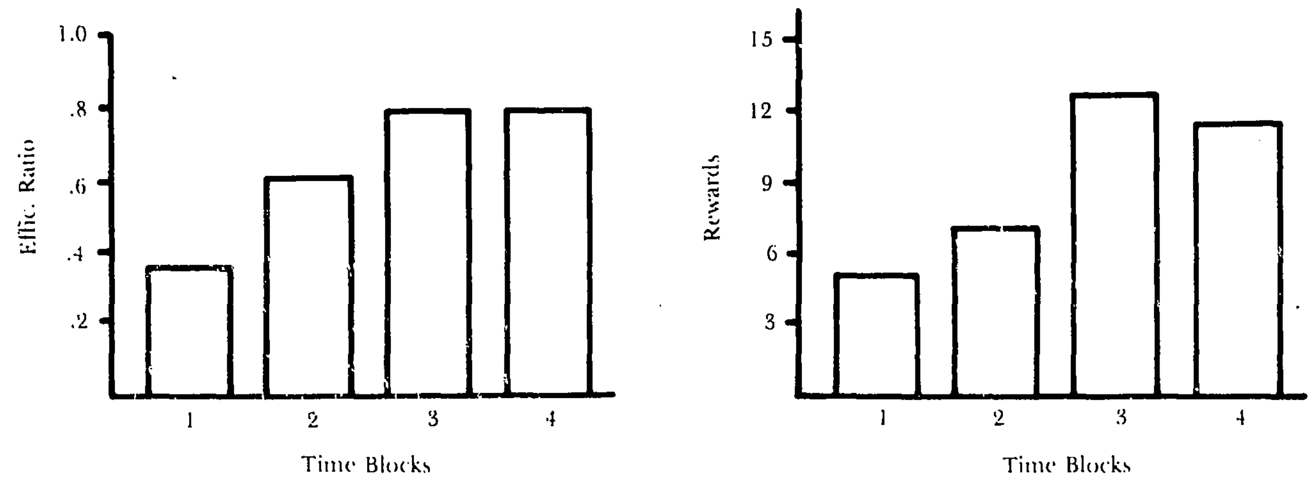
When the Total Efficiency Ratio or Total Rewards scores are abnormally low, more information about the subject can often be developed by examining the scores obtained in each of the four time blocks of the task. Examples of block by block analyses are presented below.

Profile 1



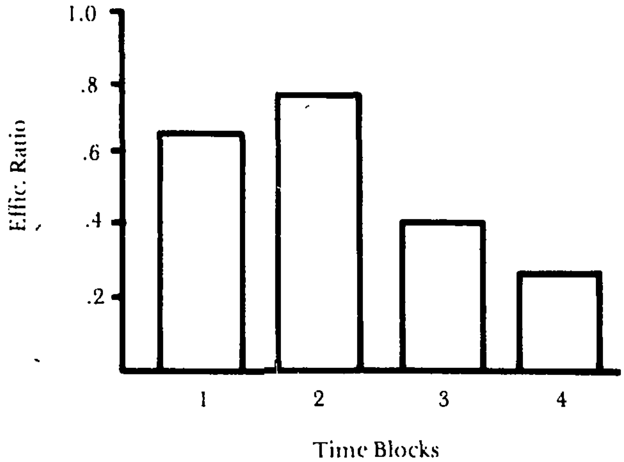
This is the profile of a youngster who demonstrates a normal overall performance on the task. We have found that such children do not significantly vary their performance from one block to another. The presence of impulsive behavior is not substantiated by these results and concern about behavior in such a child should focus on factors such as environmental stress, anxiety, or family issues.

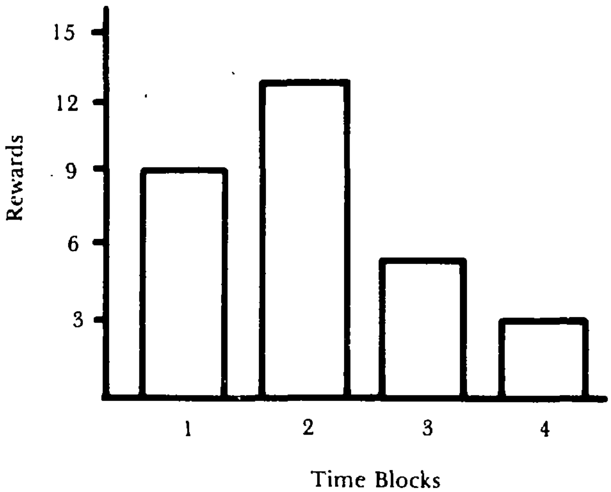
Profile 2



This profile is that of a youngster who has a difficult time initially adjusting to the demands of the task, but who can ultimately "adapt" and develop the strategies necessary to perform adequately. The relatively normal number of Rewards in the early Blocks indicates that the excessive Responses occurred in bursts that were widely enough separated to allow points to be earned. Later in the task, the child was able to suppress these bursts. Concerns about impulsivity should be addressed by emphasizing structure and considerable feedback in the early stages of any task the child is learning to perform.

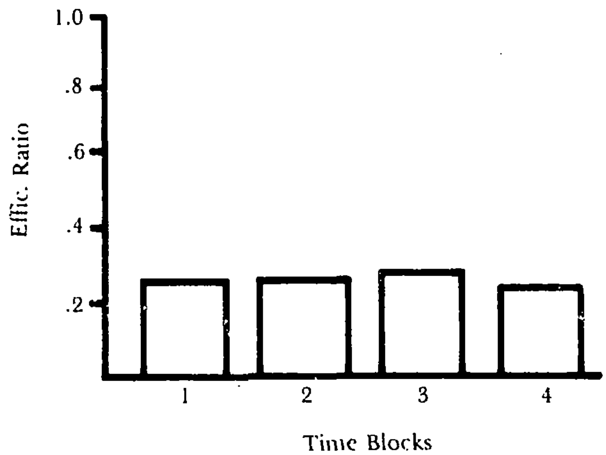
Profile 3

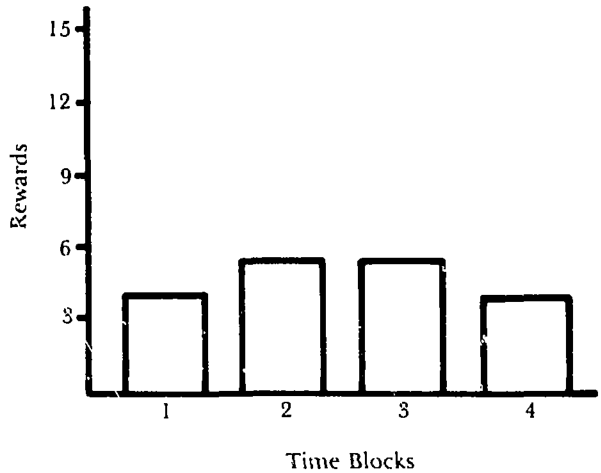




This profile is suggestive of a child who is capable of delay and suppression, but who has difficulty sustaining motivation or attention. In this case, the Rewards and Responses both diminish with time, indicating a general decrease in the child s participation in the task. Since the child is initially capable of nonhyperactive performance, issues such as environmental stress and social history may provide important information to help with the behavioral concerns.

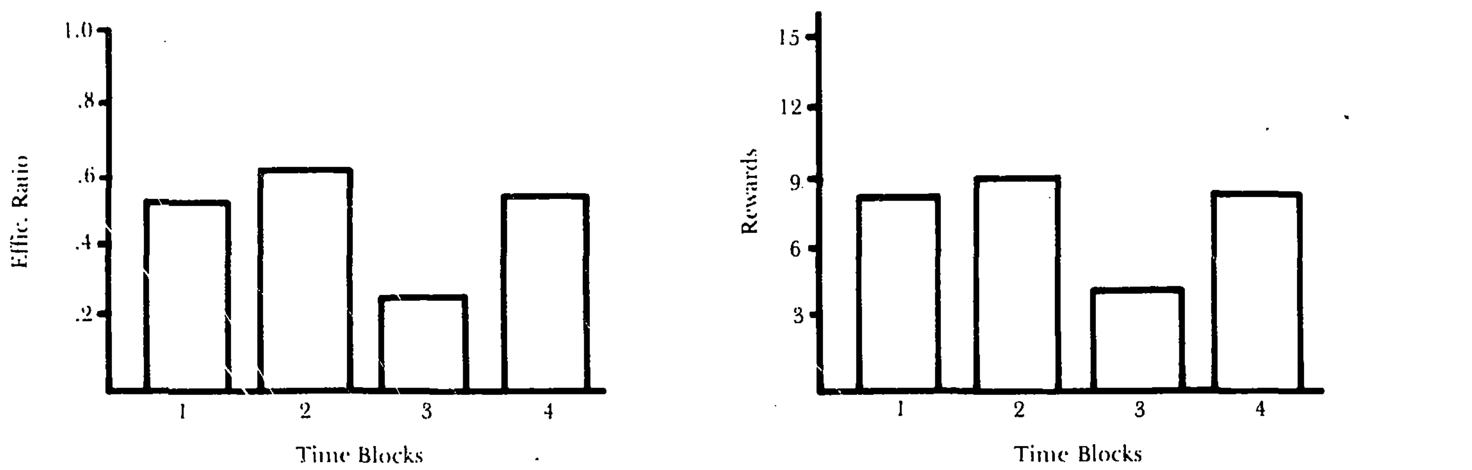
Profile 4





This pattern reflects a youngster with significant deficits in the ability to inhibit his impulses to press the button. As described under TOTAL SCORES, this behavior diminishes both the Efficiency Ratio and the Rewards. The uniform results across the time Blocks indicate that the child never mustered the necessary skills to perform well. The suspicion of impulsive behaviors is confirmed by this task performance pattern. Interventions should be geared toward cognitive-behavioral strategies and possibly stimulant medication.

Profile 5



This pattern is rather unusual, but has appeared from time to time in our clinical practice. It reflects a marginal ability to suppress responding with a rather profound lapse in the third time Block. While no definitive interpretations are offered, it is interesting to note that most children demonstrating this pattern were referred for emotional problems rather than assessment of impulsivity. This relationship will be investigated further in the future.

Evaluating Task Monitoring Data

When used in conjunction with the Data Analysis Program, the GDS provides data about the details of a child's performance during the Delay Task. The computer can display the distribution of interresponse intervals for each of the four time Blocks and for the total task. This data should be a valuable adjunct in determining why a child did poorly on the task. The child who has a low Efficiency Ratio from too many bursts of button presses will have a high frequency of response intervals in the 0-2 second range. In contrast, a child who is less impulsive may have a poor Efficiency Ratio with a high frequency of "near misses", interresponse intervals of 4-5 seconds. Research is being performed to establish interpretive standards for evaluating the interresponse distribution, but the practitioner may already be able to use the data to confirm subjective impressions about the child's difficulty with the task.

Subjective Data

As with any psychometric technique, observations of the child's behavior can be as telling as the scores themselves. Children without ADD tend to sit quietly and exhibit little actual movement throughout the session. When asked what they did to help them wait, the majority of children over 7 years of age say that they counted and that the game was easy. Younger children may display somewhat more activity than older children, but they are generally well-controlled and organized in their approach to the task. Children with an impulsive style, on the other hand, tend to have profound difficulty sitting and maintaining their motivation. They will squirm, leave their seat, run around the table, talk to the examiner and engage in a variety of other behaviors between responses. Many of these children will attempt to manipulate,the apparatus by pulling at the button, attempting to open the back panel, or by pressing on the stimulus lights. (One particularly active boy tried to throw the device down a flight of stairs.) Again, it is rare for a child without an attention deficit to exhibit such behaviors. The tasks are as uneventful for them as they are trying for there ADD counterparts.

There are behavioral patterns which are characteristic of certain groups of non-ADD children. As one might expect, unusually fearful children will tend to sit extremely still and produce very few responses. One highly anxious child from an outpatient clinic pressed the button only once in eight minutes and sat frozen for the duration. Very competitive children can become excited and challenged but will still achieve normal scores. Observations of the child's behavior during the tasks yield a bounty of information and should be recorded carefully.

GDS VIGILANCE TASK

Demands Of The Task

The GDS Vigilance Task yields data regarding the child's ability to focus attention on a task and to maintain this attention over time in the absence of reinforcement (see Description of the Task on page 4). While the primary demand of this task is maintaining attention, a child's successful performance also entails other skills as outlined below.

1. Cognitive Skills
2. Maintaining Attention
3. Behavioral Demands
4. Motivational Factors

Many of the comments already made in reference to the Delay Task also apply to the Vigilance Task. They have not been reiterated in this section so the reader is strongly advised to read the discussion of the Delay Task (pages 6 to 10).

Cognitive Skills

The Vigilance Task and its instructions are quite different from the Delay Task, but the cognitive skill requred to understand the instruction set is similar. A Correct response is registered only when the subject presses the button after a "hot 9", that is a "9" immediately preceded by a "1". These instructions require the subject to attend to two stimuli:

a. the "1" serves as the "alerting stimulus" which prepares the subject to respond

b. the "9" serves as the "target stimulus" to which the subject must respond only after having been alerted.

Unlike the Delay Task, the Vigilance Task required modification for younger children. Those under 6 years may have difficulty understanding that they are to push the button only when a "9" is immediately preceded by a "1" on the digital display. Some do not even recognize the number nine yet. The "1" mode of the Vigilance Task (page 4) is simple enough for these children. (We have found that young children perform best with a Presentation Interval of 2 seconds and a Block Lengths of 120 seconds).

Maintaining Attention

The Vigilance Task places its maximum demand upon the child's ability to maintain attention. This demand is exacting because the Task requires constant attention for nine minutes (Standard Parameters) and because no reinforcement is provided to the subject during that time. The subject's responses must be guided by his attentional and cognitive processes alone. Children who are hyperactive usually do poorly on the Vigilance Task because they cannot resist the impulse to press the button at inappropriate times, or because they miss the stimuli when distracted. Children who are not hyperactive may also do poorly on the Vigilance Task because they cannot sustain the necessary attention to press the button at appropriate times.

Behavioral Demands

In order to do well, the subject must not only respond to every "hot" target stimulus, but he must also suppress or refrain from the impulse to respond to the alerting stimulus, non-target numbers which immediately follow the alerting stimulus, and a "non-hot" target stimulus. Highly impulsive children must also resist their tendency to respond randomly. Therefore, impulse suppression is an important demand in the Vigilance Task, although the primary demands are in the area of sustaining attention.

Motivational Factors

The general motivation to succeed on this task is essentially the same as that required on the Delay Task. However, a very critical difference between the two tasks is that here the subject must pay attention for nine minutes in the absence of any external reinforcement. The apparatus provides no feedback whatsoever; the subject receives no information regarding his performance. The only feedback available is that the subject will often be aware of his errors immediately after committing them. Linked with the general motivation to do well, this internal feedback can help to sustain his attention.

Data From The GDS Vigilance Task

Outlined below are the various scores generated by the Vigilance Task:

1. Total scores
   a. Correct--responses to a "1/9" sequence
   b. Omissions--failure to respond to the "1/9" sequence
   c. Commissions--responses to other than the "1/9" sequence
2. Block Data - Correct, Omissions and Commisions for each of 3 blocks
3. Task Monitoring Data (available using the GDS Data Analysis Program) - Distribution of digit patterns preceding the responses
4. Subjective Data

Interpretation Of The Vigilance Task

The general comments made for the Delay Task (page 7) apply equally well to this task. The practitioner should evaluate all the available information before making a diagnosis and planning a therapeutic approach.

As mentioned earlier, the Vigilance Task is a version of the Continuous Performance Test developed by Rosvold and his colleagues in 1956. Although the technique has been employed in many research studies through the years, formal norms were not established until standardization studies for the GDS were completed. For technical reasons, the Vigilance Task could not be included until these studies were well underway, and our clinical experience with the task is not as extensive as that for the Delay Task. A complicating factor is that Vigilance Task scores, particularly errors of Omission, are highly age-related (see page 5), and there is a great deal of variability under the age of 8 years. Consequently, the thresholds for interpretation described below should be used with care, especially for the records of younger children. Research is in progress and should considerably increase the diagnostic power of the task.

The Vigilance Task yields two fundamental sets of results. errors of Omission, and errors of Commission. Omissions are registered when the "1/9" sequence occurs, but the subject does not respond. The alerting stimulus, the "1", did not have its intended function of preparing the child to respond to the target stimulus, "9". These errors are most frequently due to inattention, but can arise from failure to understand the instructions or insufficient motivation. (Since the total number of occurrences of the "1/9" sequence is determined only by the task Parameters, the number of Correct responses is a measure that is dependent on the number of Omissions. The number Correct is the number of "1/9" presentations, 45 in 9 minutes , minus the Omissions).

The second independent measure, Commissions, is the number of responses to any digits other than the "hot" target stimulus, i.e. a response which occurred after a number sequence other than a "9" immediately preceded by a "1". Errors of Commission fall into three basic categories:

a. redundant errors,

b. "impulsive" errors, and

c. "random" errors.

While all three types are considered to connote impulsivity, they embody different nuances of impulsive behavior.

Redundant errors are those which occur when the subject responds to the "hot 9" correctly, and then responds to the very next digit that is presented. This is considered to be a "motor overflow" error, and indicates that the subject is unable to refrain from emitting more than one response once he has been "primed" by the hot sequence.

Impulsive errors reflect any response made to the alerting or target stimulus when it is not appropriately accompanied by the other. A response to the alerting stimulus suggests that once the subject has been "primed", he is unable to suppress the response long enough to wait for the presentation of the target stimulus. A response to a "non-hot 9" similarly reflects an inability to refrain from responding.

Random errors are responses which occur to numbers other than "1", "9", or any number immediately following the "1/9" sequence. These errors are mostly likely linked to lapses in motivation or misinterpreting instructions. (Data for categorizing the errors of Commission is made available by the Data Analysis Program).

Four basic Patterns emerge in Vigilance Task performance:

1. Omissions: Normal*(< 15)
   Commissions: Normal (< 28)

This pattern reflects a subject who has sustained the attentional processes sufficiently to "catch" the expected number of "1-9" sequences, and who can also refrain from responding to non-target stimuli. These youngsters are also not impulsive.

2. Omissions: Abnormal (> 15)
   Commissions: Normal (< 28)

This pattern suggests a child whose major difficulties lies in the area of sustained attention or distractibility. There is little to suggest difficulties in the area of impulse suppression. Many of these youngsters are classified as learning disabled or as ADD without hyperactivity.

----

*Thresholds are approximate and were determined for children over 8 years of age.

3. Omissions: Normal (< 15)
   Commissions: Abnormal (> 28)

This pattern reflects a subject who can attend sufficiently in a sustained fashion, but who cannot refrain from emitting impulsive responses.

4. Omissions: Abnormal (> 15)
   Commissions: Abnormal (> 28)

This pattern is reflective of a child who not only is highly distractible and inattentive, but who is also quite impulsive. Children who present with this pattern of responses on the Vigilance Task may have some form of learning disability in addition to impulsivity which gives rise to hyperactive behaviors. They will also tend to do poorly on the Delay Task.

In summary, these two tasks provide the clinician with a wealth of information regarding a child's abilities to suppress impulsive responses and to focus and sustain attention. Our clinical work using the two tasks in tandem has been very successful, primarily because together they cover a broad range of abilities that are essential to academic performance and behavioral stability. Employing these two tasks together has enabled us to understand some situations which otherwise would have exceeded our abilities to conceptualize and treat a particular child's difficulties.

TASK INTERRELATIONSHIPS

Our research has indicated that the two GDS tasks tend to correlate in the .58 - .60 range. Children who do poorly on the Delay Task also tend to do poorly on the errors of Commission of the Vigilance Task. By the same token, children who perform well on the Delay Task tend to make few errors of Commission. Errors of Omission, on the other hand, tend to operated fairly independently of the measures geared toward the assessment of impulsivity. Thus, while the two measures tend to correlate at a relatively high level and apparently tap into some similar dimensions, in a variety of ways they also measure different abilities.

One major difference between the two task has to do with the nature of reinforcement or feedback provided by the apparatus. The provisions of feedback is an integral function of the apparatus in the Delay Mode. In the Vigilance mode the apparatus provides no feedback. The child must perform for nine minutes essentially ignorant of the success or failure of his responses. This helps to make the Vigilance Task a pure measure of attention.

From a theoretical perspective, then, it may be that the dimension of informational feedback may account, in part, for a child's performance difference between the two tasks. It could be hypothesized that children who perform well on the Delay Task and poorly on the Vigilance Task are in some ways communicating something to us regarding their ability to perform in the presence or absence of informational feedback.

CLINICAL EXAMPLES

Case 1

A seven year old girl was referred to the child development clinic by her teacher for poor school performance, not following instructions, missing assignments, and difficulty with reading. The teacher viewed much of this youngster's behavior as willful and oppositional, and was at her wit's end as to how to help the girl learn. Her parents reiterated the teachers complaints, and added that they felt the child to be distant and aloof much of the time. During the clinical interview this youngster was pleasant and cooperative, but gave the impression of either wanting to be somewhere else or of simply daydreaming.

On the Delay Task she obtained an overall Efficiency Ratio of .82, well within the normal range. She earned 46 Total rewards. Her responses appeared controlled, orderly, and goal directed. Obviously her difficulties did not lie in the area of impulse control. On the Vigilance Task her difficulties became quite clear. Although she produced just one error of Commission, which is well within normal limits, she made 33 errors of Omission, scoring well beyond the mean on this measure.

Although her motivation throughout the task was quite good, she was unable to maintain the degree of preparedness required to perform effectively on this part of the task. Therefore, we were able to identify her main difficulty as a deficit in sustained attention in the absence of impulsivity, which fits the DSM III classification of Attention Deficit Disorder without Hyperactivity.

She was placed on a very small dose of stimulant medication, and her teacher was informed of the results of the testing. Intervention strategies were offered for both home and school, and her adjustment improved in a satisfactory manner.

Case 2

The two GDS tasks used together may also be of considerable help to the clinician in identifying children who are **not** impulsive or appropriate to classify as ADD. A ten year old boy was referred to this same clinic for the evaluation of learning problems, difficulties with classmates, much "out-of-seat" behavior, and a variety of other behavior management problems. The parents shared the school's concerns, and were very eager to obtain some pharmaceutical cure for their son.

Three years prior to this evaluation, he had been diagnosed as hyperactive and had been placed on stimulant medication. His parents reported that his initial response to this regimen was good in that he became less active and less of a management problem. They reported that this "honeymoon" period was short-lived, and he quickly became grouchy, irritable, and somewhat depressed. Preferring to have the boy active and boisterous, they terminated his stimulant medication after consulting a pediatrician. However, his problematic behavior persisted, and by the time of this referral, had escalated to the point where some form of intervention was urgently indicated.

This youngster's performance on the Delay Task was solidly in the average range (.85) with 58 rewards (normal). His responses were orderly and controlled. While he appeared fidgety and restless during his task performance, he did quite well. On the Vigilance Task his scores were also within the normal range, with two errors of Omission and only one error of Commission. These results firmly suggested that the behaviors which were bringing him into conflict with his environment stemmed from factors other than attentional deficits and impulsivity. Although his behavior truly was impulsive from a descriptive standpoint, he did not demonstrate any deficit in the ability to suppress responding or sustain attention.

Following our initial evaluation, psychological testing was undertaken to identify the sources of his problem behaviors. It was found that this youngster was frequently overwhelmed by the high goals that he perceived others had for him. He was terrified of failing in school and felt he could never fulfill his parent's expectations. This gave rise to considerable anxiety, which he discharged motorically through overactivity and impulsive behavior. He was placed in outpatient psychotherapy with a family emphasis, and currently is reported to be doing quite well.

Case 3

A nine year old boy was referred with complaints of restlessness, noncomplicance, fighting with other children, and poor academic performance. While his parents had been recently divorced, these behaviors were longstanding and had been problematic since early childhood. A medical examination by his pediatrician was unremarkable, and he was referred for a psychological evaluation. Upon initial contact with the psychologist, this youngster appeared inhibited, quiet, and withdrawn, giving no indications of impulsive or hyperactive behavior. However, when confronted with the demands of the Delay Task, the boy was unable to maintain the illusion. He achieved an efficiency ratio of .44 (less than the 5th percentile) with 28 Rewards. His behavior throughout the Delay Task was disjointed, involving much out-of-seat behavior and extreme restlessness. While these behaviors in some ways helped him to suppress responding, because he was otherwise occupied, he was nonetheless unable to refrain from emitting a large number of unreinforced responses. This pattern of behavior was repeated during the Vigilance Task. He had a high number of "random errors" and consistently responded immediately after a "1". His performance was clearly suggestive of an inability to delay, particularly once he had been primed to respond. The youngster was placed on a moderate dose of stimulant medication, and was classified as having ADD with Hyperactivity. His academic programming was geared more toward accuracy than speed, and he received resource help to encourage him to modulate his response style. Follow-up contact indicated substantial improvement.

ONGOING RESEARCH

Our own research efforts concerning the clinical application of the GDS will include studies on effects of reinforcement and feedback, as well as research on the effects of stimulant drugs, patterns of responses of learning disabled children, and the possible applications of the GDS in training children to be less impulsive and more attentive. One of the major enhancements that has been developed for the GDS is an interface with a portable computer. In addition to making the total and block data more accessible, the computer has been programmed to directly receive GDS data while a task is being performed. For the Delay Task, the computer stores every interresponse interval and summarizes this data by generating a listing of the interresponse intervals. For the Vigilance Task, it saves the record of every button press and assembles a list of the frequency of digit sequences preceding the presses. This data will enable the clinician to make even more precise determinations of **why** the subject performed in certain ways.

For example, when a child makes numerous errors of Commission on the VigilanceTask, it will be very important to the clinician to know where these errors occurred in the sequence of the task. If they occur because the child tries to respond to the hot combination, but responds too late to be credited with a correct response, this is quite different from the child who responds to the "9" when it is not hot or who responds to the "1" alone. This computer innovation will have profound and exciting implications for our ability to "fine tune" our interpretations of the data provided by the GDS.

REFERENCES

Achenbach, T. The child behavior profile: I. Boys aged 6-11. *Journal of Consulting and Clinical Psychology,* 1978, 46, 478-488.

Achenbach, T. & Edelbrock, C. The classification of children's psychopathology: A reveiw and analysis of empirical efforts. *Psychological Bulletin,* 1978, 85, 1275-1301.

Buddenhagen, R. & Sickler, P. Hyperactivity: A forty-eight hour sample plus a note on etiology. *American Journal of Mental Deficiency,* 1969, 73, 580.

Cairns, E. & Cammock, T. Development of a more reliable version of the Matching Familiar Figures Test. *Developmental Psychology,* 1978, 14, 555-560.

Cohen, N., Douglas, V., & Morganstern, G. The effects of methylphenidate on attentive behavior and autonomic activity in hyperactive children. *Psychopharmacologia,* 1972, 22, 282, 294.

Conners, C. A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry,* 1969, 126, 884-888.

Douglas, V. Stop, look, and listen: The problem of sustained attention and impulse control in hyperactive children. *Canadian Journal of Behavioral Sciences,* 1972, 4, 259-282.

Egeland, B. & Weinberg, R. The Matching Familiar Figures Test: A look at its psychometric credibility. *Child Development,* 1976, 47, 483-491.

Gardner, R. *The Objective Diagnosis of Minimal Brain Dysfunction.* Cresskill, N.J.: Creative Therapeutics, 1979.

Gittis, A. & Gordon, M. Developmental analysis of behavioral dysfunction in rats with septal lesions. *Journal of Comparative and Physiological Psychology,* 1977, 94-106.

Glow, P. H. & Glow, R. A. Hyperkinetic impulse disorder: A developmental defect of motivation. *Genetic Psychology Monograph,* 1979, No. 100, 159-231.

Gordon, M. The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive boys. *Journal of Abnormal Psychology,* 1979, 7, 317-326.

Kagan, J. Reflectivity-impulsivity: The generality and dynamics of cognitive tempo. *Journal of Abnormal Psychology,* 1966, 71, 17-24.

Kendall, P. & Wilcox, L. Self-control in children: Development of a rating scale. *Journal of Consulting and Clinical Psychology,* 1979, 7, 317-326.

Maynard, R. Omaha pupils given behavior drugs. *Washington Post,* June 1970.

McClure, D. F. & Gordon M. The performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. Paper submitted to the *Journal of Abnormal Child Psychology,* 1983.

Machover, K. *Personality Projection in the Drawing of the Human Figure.* Springfield, Ill., Charles G. Thomas, 1949.

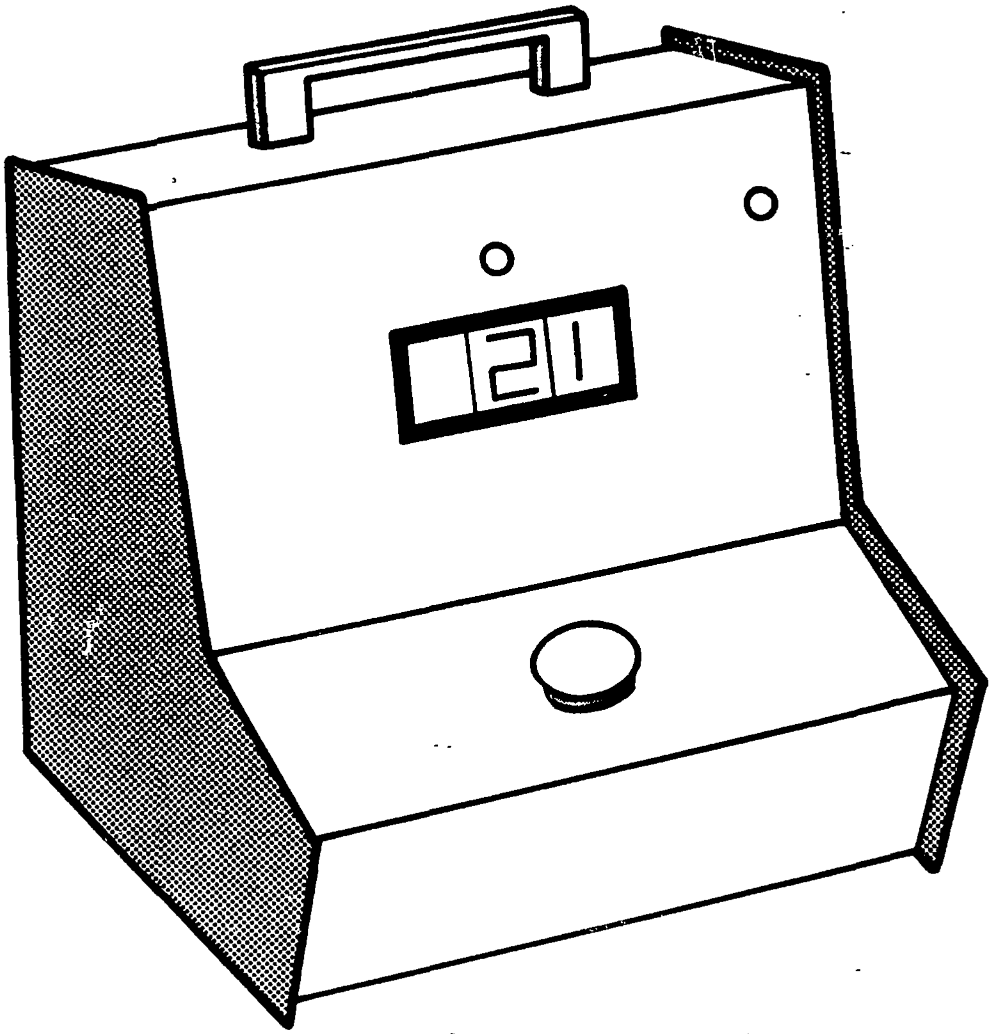
Miller, R. G., Palkes, H. S. & Stewart, M. A. Hyperactive children in suburban elementary schools, *Child Psychiatry and Human Development,* 1973, 4, 121-127.

Rosvold, H., Mirsky, A., Sarason, I., Bransone, E., & Beck. A continuous performance test of brain damage. *Journal of Consulting Psychology,* 1956, 20, 343-352.

Stone, F., Wilson, M., Spence, M., & Gibson, R. A survey of elementary school children's behavior problems. Paper presented at the annual meeting of the American Orthopsychiatric Association, New York,1969.

THE GORDON DIAGNOSTIC SYSTEM

INTERPRETIVE SUPPLEMENT



CLINICAL DIAGNOSTICS INC
1536 Cole Blvd., Golden, Colorado 80401
Suite 350 • 800-521-4503

SUPPLEMENT TO THE GDS INTERPRETIVE GUIDE

Michael Gordon, Ph.D.
Upstate Medical Center, Syracuse, New York

F. Daniel McClure, Ph.D.
D.C. Wilson Hospital, Charlottesville, Virginia

ACKNOWLEDGEMENTS

This Interpretive Supplement reflects the efforts of a talented and remarkably masochistic group. In particular, we would like to thank Ms. Barbara Bilinski Mettelman for her contributions both in coordinating our research and in managing the details of statistical analyses and other daily activities. The manuscript benefited greatly from the comments of Ms. Barbara Bass, Ms. Mary Galligan and Dr. Ernest Post from Clinical Diagnostics. Finally, we acknowledge once again the efforts of Dr. Wendy Gordon whose clinical and editorial expertise immensely enhanced this Supplement.

January 20, 1984

TABLE OF CONTENTS

INTRODUCTION

The purpose of this supplement is to provide specific guidelines for interpreting data generated by the GDS. The GDS taps into some very fundamental dimensions of functioning that are interrelated and affected by a host of other variables. The goal of the supplement is to assist the reader in gaining a thorough understanding of the Delay and Vigilance Tasks by providing a general framework for the interpretation and integration of the diverse factors inherent in GDS performance.

Some of the material from the original Interpretive Guide is repeated here for purposes of clarity and elaboration. We also want to point out that the logic underlying this Supplement served as the basis for the Data Analysis Program (DAP) we developed for use with microcomputers. The DAP generates the kinds of interpretations described in these pages but at a much faster rate.

As with other psychometric data, information from the GDS cannot be viewed in isolation; rather, it must be considered within the overall context of data from clinical observation, social/developmental history, and other psychodiagnostic procedures. In addition, individual GDS scores should only be interpreted in relationship to the other GDS scores and to the subject's overall performance on the tasks. The reader will find, however, that interpretation of GDS performance is straightforward and can be achieved with relative ease because it flows logically from a common-sense understanding of the tasks. It should be kept in mind that the goal of our interpretation is not only to establish normality or abnormality but also to understand how the child functions and translate that information into specific recommendations for treatment.

In this supplement we will first analyze the interpretive significance of individual scores from each test separately. Critical parameters will be identified and explained. We will then explore how the Task Monitoring data clarifies the data from each task. Finally, we will outline a method for integrating the data from both tasks into a sound and coherent overall interpretation of a subject's functioning. As we go through this process we will include interpretive statements that will help the reader understand how the authors describe and organize GDS data. These statements are not intended to be taken as cookbook interpretations. Rather, their purpose is to illustrate our thinking in a practical and concrete way.

Keep in mind that all interpretations are based upon administration of the tasks according to the Standard Parameters. It should also be noted that we do not discuss the "1" mode of the Vigilance Task. A separate supplement will cover the testing of very young children.

THRESHOLDS FOR INTERPRETATION

The first question to be answered for any GDS data is whether a particular score is normal or abnormal. Based upon data from standardization projects, we have divided scores into three ranges. A score is considered ABNORMAL if it is typical of less than 5 percent of the normal population, i.e., the 5th percentile or less. Scores in the BORDERLINE range are those which fall between the 6th and 25th percentile. Finally, scores that are above the 25th percentile are classified as NORMAL. These cutoffs correspond closely to clinical experience and our research studies in that children classified as hyperactive by other measures typically perform in the ABNORMAL range of GDS scores.

To facilitate the interpretive process, we have constructed Threshold Tables which list the ranges of scores corresponding to each category (ABNORMAL, BORDERLINE, and NORMAL). These tables are located at the end of the Supplement, and are reproduced in each section as needed. They are based on the same data as the standardization tables displayed at the end of the Interpretive Guide.

Because the Delay Task norms do not vary significantly with age, we have presented one table for all children from 6 to 12 years. Since some of the Vigilance Task variables are age-related, expanded tables were required. Comparison of an individual child's scores to these tables will determine the degree of normality or abnormality.

INTERPRETATION OF THE DELAY TASK

The GDS Delay Task yields three kinds of Summary Scores: the number of Responses (R); the number of Rewards or correct responses (REW); and the Efficiency Ratio (ER), which is a ratio obtained by dividing REW by R. For the purposes of interpretation, we pay most attention to the ER for the whole session – the Total ER.

The Efficiency Ratio (ER)

NORMAL	.77 – 1.0
BORDERLINE	.55 – .76
ABNORMAL	⩽.54

The ER is the basic building block of Delay Task interpretation. It represents the percentage of times the child pressed the button after having waited the appropriate number of seconds. The ER is the best single indicator of the level of impulsivity demonstrated by a subject. According to the Threshold Tables, an ER of .54 or less falls in the ABNORMAL range and indicates a significant deficit in the ability to inhibit responding. The chances are that a child with an ABNORMAL ER tends to display poor self-control and would be considered by most parents and teachers to be impulsive. Children with an ABNORMAL Total ER experience difficulty in situations which require them to come up with ways to keep themselves from engaging in activity. They are overly disinhibited and have not developed strategies to help them delay when delaying would serve them well. By the same token, scores between .55 and .76 are in the BORDERLINE range, and suggest tendencies toward impulsive responding. Obviously, the closer the ER approaches the lower end of the BORDERLINE range, the stronger the tendencies toward impulsivity.

Listed below are the kinds of interpretations and recommendations that can be formulated based on the Total ER:

Interpretive Statements

Total ER within the NORMAL range:

Subject demonstrates normal capacity for delay and inhibition. If this child were referred for impulsive behavior, a more extensive psychological evaluation should be considered because problem behaviors do not appear to stem from a fundamental deficit in impulse inhibition. Anxiety, emotional difficulties, and/or environmental stress may account for these behaviors.

Total ER within the BORDERLINE range:

Subject demonstrates tendencies toward impulsive responding. S/he will require increased external structure in the classroom, particularly when confronted with tasks which involve setting one's own pace and inhibiting behavior. This child might also benefit from special programs designed to teach strategies for developing cognitive self-control over behavior. In other words, this child will likely require help in learning how to "stop, look, listen, and think" before acting. Stimulant medication is not usually the treatment of choice for such a youngster, although in some cases a brief period of pharmacotherapy may facilitate other interventions.

Total ER within the ABNORMAL range:

Subject demonstrates significant deficits in areas related to the ability to delay. S/he will require greatly increased external structure in the classroom, particularly when confronted with tasks which involve self-control and the capacity to inhibit responding. This child would benefit from special programs designed to teach strategies for developing cognitive self-control over behavior. This child will require help in learning how to "stop, look, listen, and think" before acting. Treatment with stimulant medication may be a viable option to consider as part of a general program of educational and psychological management.

While the Total ER can provide an overall pegging of a child along the continuum of impulsivity, the fuller meaning of this score can be extracted by a careful analysis of the Block Scores and Task Monitoring Data (see below).

Total Number of Responses (R)

NORMAL	21 – 56
BORDERLINE	57 – 66
ABNORMAL	⩽20 & ⩾67

Total R reflects the subject's overall rate of responding. The primary significance of R is that it determines the validity of the Efficiency Ratio. A Total ER of .80 based on only 10 responses must be interpreted with caution because the child is not attempting to gain points at a rate typical of peers. Our experience indicates that children with low R are often fearful and overly-cautious. It is typical that further evaluation shows them to experience emotional/motivational problems. It can also occur that very impulsive children refrain from responding at all for fear that any response would lead to a torrent of poor control.

Very high levels of R (especially 90 or more) also take on significance. These children exhibit a marked inability to suppress responding. They may be able to score Rewards but there is a flood of impulsive responding. Of course, most children with a high number of Responses will also have poor ER's and will be considered impulsive. When very high responsivity occurs, however, it is important to look for patterns of impulsivity (see below). In addition, the subject's motivation to perform well on the task should be considered suspect.

Total Number of Rewards (REW)

NORMAL	38 – 80
BORDERLINE	22 – 37
ABNORMAL	⩽ 21

Our data regarding the discriminative validity of the REW score has been equivocal. In one of our research studies, this variable significantly differentiated "hyperactive" from "nonhyperactive" groups. However, in another study REW failed to differentiate the experimental group from the control group. We tend not to interpret REW by itself, primarily because its impact is subsumed under interpretation of the ER.

The Analysis Of Block Scores

The GDS is internally programmed to separate the 8-minute trial into four 2-minute time blocks. This was engineered into the GDS because patterns of responding over time can be critical in the evaluation of attention and inhibition. Block Scores enable the practitioner to view how a given subject's performance has progressed over the course of the eight minutes and to detect changes in the pattern of responding. To accomplish this, the GDS yields values for R and REW for each two minute block. The resultant ER values are subscripted with the appropriate time block. For example the ER derived from Time Block 1 is notated as ER_1, the ER for Time Block 2, ER_2, and so on.

Data from both normal and clinical populations has demonstrated the GDS to have a remarkable degree of internal consistency, whether the subject is impulsive or non-impulsive. Performance is particularly consistent after the first block, during which time the subject adapts to the task. Therefore, marked inconsistency in a child's scores from block to block takes on clinical significance even for subjects who demonstrate the capacity to perform within the normal range on all variables.

To analyze Block Stores we consider 3 parameters: overall variability, slope, and extreme values.

Block Variability

NORMAL	0 – .10
BORDERLINE	.26 – .11
ABNORMAL	⩾.27

Calculations for Block Variability are aimed at characterizing the overall degree of consistency across the session. For the purposes of statistical analysis and our Data Analysis Program, the standard deviation of the four Block ER scores is used as an indicator of variability. To compute the standard deviation precisely you would follow these steps:

1. Take the Mean (average) of the 4 ER Block Scores by adding them together and dividing by 4.
2. Subtract the Mean from each Block ER.
3. Square each difference and add the results of the 4 squares together.
4. Divide by 3 and then take the square root of the quotient.

To illustrate, a subject with Block ER's of .45, .67, .44 and .78 would have a mean ER of .59. Subtracting the mean from each ER and adding the square of each difference together yields a sum of .08. Dividing by 3 and taking the square root gives us a Block Variability score of .17. We can then compare that score to the Threshold Tables to judge overall consistency*.

A child with NORMAL Block Variability will be relatively consistent in his/her efforts to delay, regardless of whether those efforts are successful or not. Children with BORDERLINE or ABNORMAL scores will be less consistent.

Interpretive Statements

Block Variability NORMAL:

Subject demonstrates normal variability in performance, indicating that s/he will be no more inconsistent in delay or inhibition than peers.

*A simpler method for judging variability entails determining the Mean ER score (by adding the Block ER scores together and dividing by 4). If any Block ER falls above or below the Mean ER by more than .15 then there is a significant amount of variability. In the example given above there is significant variability in that one Block ER falls above or below the Mean ER of .59 by more than .15 (that is, the scores fall outside the bracket of .45 to .74).

Block Variability BORDERLINE:

Subject demonstrates a degree of variability which is in the borderline range. S/he is likely to be more inconsistent than peers in the level of impulsivity or inhibition demonstrated. Interventions should focus on the development of strategies geared toward achieving stability of responding. These might involve "talking" one's self through rote tasks, practicing step-by-step sequential tasks, and outlining on paper the various steps involved in solving a particular problem.

Block Variability ABNORMAL:

Subject demonstrates a significant degree of variability in his/her performance. S/he will be unpredictable in the level of impulsivity exhibited. As with those subjects whose variability is in the BORDERLINE range, interventions should be geared toward achieving and maintaining a stable, consistent response style. Such approaches might involve "talking" one's self through rote tasks, practicing step-by-step sequential tasks, and outlining on paper the various steps involved in solving a particular problem.

Slope

An important aspect of our interpretation of Block Scores centers around discerning patterns of variability. While the Block Variability score characterizes the general level of consistency, the Slope score determines if there were any clear trends from block to block. In other words, did the subject's scores progressively improve, stay the same, or deteriorate across the session? Again, most children's ER's are relatively stable from block to block. Inspection of the ER's, particularly when graphed, gives a sense of whether or not there was improvement or deterioration. For statistical purposes we use the following formula:

$$\text{Slope Score} = (ER_3 + ER_4)/2 - (ER_1 + ER_2)/2$$

If the Slope Score is between .12 and -.12, performance is characterized as having an Even Slope. Slope Scores of more than .12 reflect an Upward Slope (i.e., improvement) while scores below -.12 reflect a Downward Slope (i.e., Deterioration).

Interpretive Statements

Even Slope:

Subject demonstrates consistent performance across blocks.

Upward Slope:

Subject demonstrates initial tendencies toward impulsivity, but can adjust to the task and achieve normal responding. After the initial difficulty in adaptation, efficient performance is achieved. This child will require more structure and support in the beginning stages of a task or exercise than his/her peers.

Downward Slope:

Subject demonstrates a tendency toward deterioration in performance over time. May require additional structure to support responding as tasks progress temporally.

Analysis of extreme scores: Peaks and Valleys

Another important factor involves any Block Score that stands in contradistinction to the Total ER. For example, a subject with an ABNORMAL Total ER (below .55), but with at least one Block ER in the normal range (i.e., "a peak") is demonstrating some capacity for normal delay. In contrast is the subject with a similar ER but with no Block ER in the normal range (i.e., above .74). This subject at no point in the testing evidences a capacity to delay adequately. By the same token, a subject with a NORMAL Total ER, but with one Block ER below .55 (i.e., "a valley") is demonstrating a potential for impulsivity under certain conditions, while a subject with a normal Total ER and all normal Block ER's is consistently well-controlled.

Interpretive Statements

Total ER BORDERLINE OR ABNORMAL and any Block ER greater than .74:

This subject does have the capacity for normal delay.

Total ER BORDERLINE OR ABNORMAL and no Block ER greater than .74:

Performance never reaches the normal range.

Total ER NORMAL and 1 Block ER ABNORMAL (less than .54):

Despite overall normal performance, subject evidences some potential for impulsive responding.

Examples of Delay Task Interpretation

To clarify issues presented in the previous sections, let us look at some sample Delay Task patterns.

SUBJECT 1 – 8 year-old male

BLOCK	1	2	3	4	TOTAL
REWARDS	13	14	13	15	55
RESPONSES	17	16	16	17	66
EFFICIENCY RATIO	.77	.87	.80	.88	.84

INTERPRETATION: This subject's Number of Responses (R) indicates that these results may be interpreted with confidence, as responsivity is within normal limits. The Total ER indicates that he has no difficulties with delay. This ability to inhibit responding persists across time, suggesting that it is a relatively stable trait. This child does not appear to have an impulsive style, *per se.* We would suggest that the practitioner look toward other areas to account for the problem behaviors.

SUBJECT 2 – 10 year-old male

BLOCK	1	2	3	4	TOTAL
REWARDS	10	9	12	13	44
RESPONSES	27	19	23	22	91
EFFICIENCY RATIO	.37	.47	.52	.59	.48

INTERPRETATION: This subject's rate of responding (R) is abnormally high so that the results should be interpreted with caution. When the number of Responses is above 90, it indicates an extreme inability to suppress responding which can be related to motivational issues. A more thorough psychological evaluation would likely be suggested to evaluate the various components of the child's difficulties with inhibition.

This subject does indeed demonstrate significant impulsivity and inability to delay. While he improves in this area over time, at no time do his Block ER's approach the NORMAL range. He will require extra structure particularly at the initial stages of a task. He would also benefit from a training program aimed at developing strategies for self-control.

SUBJECT 3 – 9 year-old female

BLOCK	1	2	3	4	TOTAL
REWARDS	12	10	6	8	36
RESPONSES	15	17	15	19	66
EFFICIENCY RATIO	.80	.59	.40	.42	.54

INTERPRETATION: This subject's rate of responding indicates that these results may be interpreted with confidence, as responding is well within normal limits. She demonstrates a significant degree of impulsivity. However, she also demonstrates the ability to delay and inhibit in a normal manner in the initial stages of a task. This ability tends to deteriorate over time, suggesting that she will require additional structure as tasks progress.

Task Monitoring Data

Task Monitoring capability was programmed into the GDS in an effort to provide the user with a more detailed view of the subject's response style. Task Monitoring Data can only be collected with the GDS Data Analysis Program. While there are not yet enough normative data to make specific interpretive statements associated with Task Monitoring Data, they allow the user to "fine tune" interpretations of the other data generated by the Delay Task. The Task Monitoring Data provide a detailed summary of the response latency from the previous response for each response emitted. This data is provided for each 2-minute time block as well as for the entire test. An example is provided by the next subject.

SUBJECT 4 – 8 year-old boy

INITIAL RESPONSE LATENCY = 0.4 SEC.

INTERVAL	BLOCK 1	2	3	4	TOTAL
0 SEC	0	0	0	0	0
1 SEC	0	1	0	0	1
2 SEC	0	2	1	1	4
3 SEC	0	0	3	0	3
4 SEC	5	1	4	5	15
5 SEC	3	4	4	2	13
6 SEC	2	0	5	4	11
7 SEC	0	1	2	2	5
8 SEC	4	1	1	3	9
9 SEC	2	2	1	2	7
10 SEC	0	2	1	0	3
11 SEC	0	0	0	0	0
12 SEC	0	2	0	0	2
13 SEC	1	0	0	0	1
14 SEC	0	0	0	0	0
15 SEC	0	0	0	0	0
16 SEC	0	0	0	0	0
17 SEC	0	0	0	0	0
18 SEC	0	0	0	0	0
19 SEC	0	0	0	0	0
>19 SEC	0	0	0	0	0

In this case the Task Monitoring Data proved to be extremely helpful in analyzing data from the Delay Task. This youngster's other scores were as follows:

BLOCK	1	2	3	4	TOTAL
REWARDS	10*	8	10	11	39*
RESPONSES	18	16	22	19	75
EFFICIENCY RATIO	.55	.50	.45	.57	.52

*The first reward in Block 1 is given regardless of the initial latency and does not appear on the table of response intervals.

This protocol is valid (as indicated by R), and the subject has significant deficits in the ability to inhibit and delay responding (as indicated by ER). Moreover, the low degree of variability suggests that this trait is fairly stable, and Block ER's indicate that this subject was at no time able to achieve and maintain a normal level of responding. However, his Task Monitoring Data sheds a great deal of light on these scores. Analysis of this data indicates that 39 of his 75 responses occurred within 4-6 seconds of the previous response, while only 8 occurred within 0-3 seconds of the previous response. More than half of his responses clustered very closely to the 6-second criterion. This suggests that, although he did quite poorly on the test, he was indeed making some relatively stable attempts, albeit unsuccessful, to delay. His problem appears to be in arriving at a strategy that would help him delay the extra 1 - 2 seconds necessary to gain a point. For some reason, he did not undergo the process of "bracketing" or testing his responses so that they would be successful. It may be that this is a child who does not sufficiently incorporate feedback to correct and improve performance. In other words, something "got in the way" of his ability to use the unrewarded responses to adjust his strategy. We have found that some children who present with this protocol become so preoccupied and anxious about "winning" that they fail to settle down to a comfortable and productive session. For this reason, we often recommend that additional psychological testing be conducted to evaluate motivational/emotional factors.

This child's Task Monitoring Data is in contrast to that of intensely impulsive children who generate profiles marked by responses occurring between 1 or 2 seconds of the previous response. These are the more typically impulsive children who exhibit no resources for developing a strategy or for delaying at all.

Thus, the Task Monitoring Data can prove to be quite valuable in refining the overall interpretation. The information provided by the TMD not only helps in the diagnostic process, but can also provide critical input into the development of remedial and educational strategies.

INTERPRETATION OF THE VIGILANCE TASK

The Vigilance Task yields three kinds of Summary Scores: the number of Correct Responses (CR), the number of Errors of Omission (EO), and the number of Errors of Commission (EC).

Correct Responses (CR)

	6–7 YEAR OLDS	8–9 YEAR OLDS	10–11 YEAR OLDS
NORMAL	⩾23	⩾37	⩾41
BORDERLINE	7–22	28–36	38–40
ABNORMAL	0–6	0–27	0–37

This score reflects the total number of "1/9" combinations that were responded to by the subject. The sequence of digits presented is fixed and there are 15 "1/9" presentations in each 3-minute time block, for an overall total of 45 possible CR for the 9-minute session. The CR score reflects the subject's level of "vigilance" or alertness. In this respect, the term vigilance refers to the ability to focus the attentional processes in a goal-directed manner and to maintain this investment of attention over time. The CR score is an index of the subject's ability to achieve and maintain alert, vigilant responding. Children with CR scores in the normal range are thus able to maintain vigilance and attention appropriately, while children with BORDERLINE and ABNORMAL CR scores are those who tend to experience deficits in sustaining attention. It should be pointed out that a CR score within the ABNORMAL range, especially as unusually low as 4 or 5, may be due to the child's lack of motivation to maintain alertness. For this reason, we often suggest that these children be closely evaluated for emotional/motivational difficulties.

Interpretive Statements

Correct Responses within NORMAL range:

Subject is capable of investing sufficient energy in the attentional processes and maintaining this investment over time. S/he demonstrates intact vigilance under conditions of relatively high arousal. If this child was referred for inattentiveness, a more extensive psychological evaluation should be considered because problem behaviors do not appear to stem from a fundamental deficit in sustained attention. Anxiety, emotional difficulties, and/or environmental stress may account for these behaviors.

Correct Responses within BORDERLINE RANGE:

This subject's vigilance and alertness fall within the BORDERLINE range. Lapses in sustained attention are evident in his/her performance. Remedial strategies should involve helping him/her to remain attentive and to discriminate salient from non-salient stimuli. S/he may need instructions simplified and repeated. During task performance this youngster may require more supervision and monitoring than peers. Pharmacotherapy may be helpful as an augmentative intervention.

Correct Responses within ABNORMAL range:

This subject demonstrates significant deficits in sustained attention, which are possibly related to emotional or motivational difficulties. S/he is not alert to important cues, and may frequently miss or overlook salient aspects of a stimulus situation. This child might benefit from the "cognitive focusing" effect that has been noted with regimens of stimulant medication. Augmentative interventions should also include exercises geared toward learning attentiveness in high-arousal situations, such as flash-card exercises.

Errors of Omission (EO)

This score reflects the number of "1/9" presentations to which no response was made. Since the CR score reflects the number of these combinations to which the subject responded, the CR and EO scores mirror each other. For example, if a subject obtained a CR score of 41, then his or her EO score would have to be 4 (41 correct out of a possible 45). Consequently, the EO and CR scores are two sides of the same coin and receive identical interpretations. As with the CR score, the EO score is an index of lapses in this alert, vigilant attitude. These errors reflect instances in which the subject was not attending to the situation sufficiently to respond to the "1/9" sequence.

Errors of Commission (EC)

	6–7 YEAR OLDS	8–9 YEAR OLDS	10–11 YEAR OLDS
NORMAL	0–14	0–11	0–6
BORDERLINE	15–49	12–29	7–16
ABNORMAL	⩾50	⩾30	⩾17

The EC score reflects the number of times the subject responded to incorrect stimuli, that is, stimuli other than a

"9" immediately preceded by a "1". These errors are largely unrelated to either Correct Responses (CR) or Errors of Omission (EO), and actually constitute data on a separate dimension. While CR and EO reflect the adequacy of sustained attentional capabilities, EC reflect the degree of impulsivity the subject exhibits under the structure imposed by the cognitive demands of this task. The EC represents an index of the subject's inappropriate responding; i.e., responses which reflect not lapses in sustained attention, but rather responses to stimuli in an ill-conceived or poorly controlled fashion. While we will be discussing different categories of EC which reflect various aspects of a subject's functioning, keep in mind that all EC represent impulsive responding in one way or another.

Interpretive Statements

EC within NORMAL range:

Under conditions demanding vigilance and alertness, this subject demonstrates a level of impulsivity which is no different from his/her peers.

EC within BORDERLINE range:

Under conditions demanding vigilance and alertness, this subject demonstrates a level of impulsivity which is somewhat greater than his/her peers. S/he is likely to become disorganized during competitive classroom exercises, particularly when speed of responding is a requisite. Activities such as these should be avoided with this child until interventions geared toward reducing impulsivity under these circumstances can be effected.

EC within ABNORMAL range:

Under conditions demanding alertness, this subject demonstrates a high level of impulsivity. S/he is very likely to become disorganized during competitive classroom exercises, particularly when speed of responding is required. Competitive activities should be avoided. Evaluation for pharmacotherapy should be considered. As with many other very extreme GDS scores, these scores may reflect, in addition to a possible impulsive style, difficulties in motivation, lack of comprehension of the instructions, or interference by emotional factors. Further psychological evaluation is often recommended.

Errors of Commission fall into two basic categories which can be determined by inspection of the Task Monitoring Data:

Target-related Errors: These are the most frequent Errors of Commission. A Target-related Error is registered by any response to either the target stimulus ("9") or to the alerting stimulus ("1") when they are not in the correct "1/9" sequence. For example, a response to the alerting stimulus prior to the presentation of the next digit is a Target-related error. By the same token, a response to the target stimulus when it is not immediately preceded by the alerting stimulus also registers a Target-related error. A response to the "1" before the presentation of the next digit suggests that the subject is unable to inhibit responding long enough to await the next stimulus and scan it for correctness. A response to any non-target stimulus which immediately follows a "1" reflects that once the subject has been "cued" or primed to respond by the alerting stimulus he or she is unable to refrain from emitting a response, even if it is incorrect to do so. Similarly, responses to the "9", when not immediately preceded by the "1", reflect a disregard for the salient aspects of the stimulus situation.

Sometimes the subject responds correctly to the "1/9" sequence, but then responds to the very next digit presented, whatever that particular number may be. These errors reflect the inability to inhibit a response immediately after emitting a correct response. Most often, they relate to an elevated level of arousal which cannot be brought under control sufficiently to prevent an immediate, impulsive second response.

Random Errors: A Random Error is any button press that does not occur in response to either the "alerting stimulus" or to the "target stimulus". Thus, when the subject presses the button in response to a "4" or to a "7", when these stimuli were not immediately preceded by the "1" or "9", a Random Error is registered. These errors are highly unusual, and generally reflect a state of arousal which is out of control. However, a caveat applies here. In a protocol with many of these responses, the subject's motivation to perform well and comprehension of the instructions must be examined.

Block Scores

Just as with the Delay Task, the Block Scores reflect the overall variability in a subject's performance. This variability should be interpreted in the context of the total scores. For example, if the Total CR and EC are both NORMAL, the significant variability takes on a different meaning than when the Total Scores are BORDERLINE or ABNORMAL.

For the Vigilance Task Block Scores we examine two parameters, namely Slope and Extreme Scores (Peaks and Valleys):

Slope

As with the Delay Task we look for trends or patterns across the three blocks of the session. Most children maintain a consistent performance (Even Slope), but some improve (Upward Slope) and a few deteriorate (Downward Slope). These patterns can be estimated by visual inspection of the scores. For statistical purposes, we define a protocol as having an Upward Slope when each block has a greater number of responses (Correct or Commission) than the previous one. Conversely, a Downward Slope is when each block has fewer responses (Correct or Commission) than the preceding block. Keep in mind that an Upward Slope for CR reflects improvement while for EC it indicates deterioration.

Interpretive Statements

Even Slope for CR:

Subject demonstrates consistent performance across time.

Upward Slope for CR:

Subject initially demonstrates lapses in sustained attention, but performance improves over time as s/he becomes more familiar with the task.

Downward Slope for CR:

This subject demonstrates an initial ability to sustain attention and respond adequately, but his/her performance deteriorates over time.

Even Slope for EC:

This subject demonstrates consistent performance with respect to behavioral suppression.

Downward Slope for EC:

Performance improves over time with respect to behavioral suppression. In early stages of a task, performance may be disorganized and haphazard but becomes better controlled as subject adjusts to task demands.

Upward Slope for EC:

Performance deteriorates with respect to behavioral suppression. Subject becomes less controlled as task progresses.

Peaks and Valleys

In a protocol which is significant for deficits in sustained attention and/or behavioral suppression (CR and/or EC in ABNORMAL range), any Block Score along these two dimensions which is within normal limits must be considered in the interpretation.

Interpretive Statements

Total CR in BORDERLINE or ABNORMAL range and any Block CR greater than 8 for 6 & 7 year-olds and greater than 11 for the older groups:

While this subject demonstrates overall deficiencies in sustained attention, there are times when s/he is capable of sustaining attention as well as his/her peers. However, this ability is not of sufficient strength to bring overall functioning within the normal range.

Total CR in NORMAL range and any Block CR less than 4 for 6 & 7 year-olds and less than 9 for the older groups:

While this subject demonstrates essentially normal capacities for sustained attention, there are times when his/her attentional processes will falter to a level below that of his/her peers.

Total EC in BORDERLINE or ABNORMAL range and any Block EC less than 5:

Subject demonstrates capacity for normal behavioral suppression.

Total EC in NORMAL range and any Block EC greater than 10:

While subject demonstrates generally normal behavioral suppression, there are times when s/he will become impulsive when required to sustain attention.

RELATIONSHIP OF THE EC SCORE TO DELAY TASK SCORES

Interpretations of impulsivity based solely on Vigilance Task performance must be made with extreme caution. The GDS Vigilance Task places specific cognitive demands upon the subject, and the emergence of impulsive responding under these conditions may not necessarily reflect a core dimension of impulsivity. The Vigilance Task may stand alone when addressing issues related to the attentional processes and the ability to sustain attention. However, the interpretation of EC scores rests heavily upon the subject's performance on the Delay Task. The Delay Task taps into a central dimension of impulsivity, while EC scores reflect impulsivity which emerges or is suppressed under certain conditions.

Examples of Vigilance Task Interpretation

Three profiles are presented below to illustrate interpretation of the Vigilance Task. The Task Monitoring Data are displayed in a table that lists the frequency with which the subject responded to the various combinations of digits. In this table an "X" represents any digit that may have appeared. For example, the "X19" sequence indicates that the button was pressed upon presentation of a "9" which had been preceded by a "1", which in turn had been preceded by any number ("X"). This row would therefore show the number of Correct Responses. The "XX1" sequence records the number of times the button was pressed upon presentation of a "1" preceded by any two other numbers.

SUBJECT 5 – 9 year-old male

BLOCK	1	2	3	TOTAL
CORRECT	13	10	7	30
OMISSIONS	2	5	8	15
COMMISSIONS	14	23	37	74

SEQUENCE BLOCK:	1	2	3	TOTAL
X19	13	10	7	30
XX1	1	2	9	12
19X	1	2	3	6
XX9	0	0	0	0
X9X	2	6	4	12
X1X	7	4	6	17
XXX	3	9	15	27

This youngster evidences many of the hallmarks of an impulsive style. The Total CR of 30 is in the ABNORMAL range and indicates significant deficits in his ability to sustain attention in a goal-directed manner. He is apt to be quite inattentive in situations which require vigilance, such as a classroom setting. Block Scores indicate that, while his initial investment of attention is adequate, it deteriorates steadily as the task progresses. The EC score also suggests an extreme level of impulsive responding in situations requiring vigilance. As his capacity for sustained attention diminishes, his level of impulsive responding increases.

The Task Monitoring Data indicates that the majority of this child's Errors of Commission were "Target-related". This suggests that much of his impulsivity is related to his level of arousal: when he is "primed" to respond by the target stimulus he is unable to inhibit sufficiently to insure the appropriateness of his response. The Task Monitoring Data also reveal another important aspect of his performance. Note that in the first Time Block there were only 3 Random Errors (see the "XXX" row on the table); most of his errors were "Target-related". As the task progressed, his level of arbitrary responding (i.e., Random Errors) increased dramatically. This suggests that, as he becomes less able to sustain attention, he becomes more arbitrary and uncontrolled. It would be important to examine his Delay Task performance to determine his level of impulsivity in a self-paced task where arousal level becomes less critical.

SUBJECT 6 – 11 year-old boy

BLOCK	1	2	3	TOTAL
CORRECT	6	8	12	26
OMISSIONS	9	7	3	19
COMMISSIONS	16	8	5	29

SEQUENCE BLOCK:	1	2	3	TOTAL
X19	6	8	12	26
XX1	9	1	3	13
19X	6	0	1	7
XX9	0	0	0	0
X9X	0	0	0	0
X1X	1	7	1	9
XXX	0	0	0	0

This subject demonstrates deficits in sustained attention as evidenced by his ABNORMAL CR score of 26. He will therefore be apt to miss salient aspects of a situation. The Block Scores indicate that his inattentiveness and his impulsivity, both of which were problematical initially, improved over time. As such, this youngster will tend to be disorganized and inattentive upon the initial presentation of a task but can consolidate his performance as he "settles in".

A review of the Task Monitoring Data indicates that, while he is impulsive in the early stages of the task, his responding never became arbitrary (i.e., there were no Random Errors). The bulk of his errors were Target-related suggesting that once primed to respond, he is relatively unable to modulate his arousal in order to suppress inappropriate responses. He is thus likely to become over-stimulated when faced with competitive tasks or when speed is a major task requirement. Examination of his Delay Task performance would clarify the degree to which this impulsivity is manifested in a low-arousal situation.

SUBJECT 7 – 7 year-old girl

BLOCK	1	2	3	TOTAL
CORRECT	13	15	13	41
OMISSIONS	2	0	2	4
COMMISSIONS	2	1	0	3

SEQUENCE BLOCK:	1	2	3	TOTAL
X19	12	15	13	41
XX1	0	0	0	0
19X	0	0	0	0
XX9	0	0	0	0
X9X	0	1	0	1
X1X	2	0	0	2
XXX	0	0	0	0

This child demonstrates unusually well-developed capacities for sustained attention and behavioral suppression in situations requiring vigilance. Such a "clean" protocol is unusual for a youngster of her age, particularly since she was referred for an evaluation because of inattentiveness and poor self-control. Her Delay Task scores were also well within the NORMAL range. It was therefore suggested that the practitioner look toward motivational/emotional or environmental factors to account for the presenting complaints.

INTEGRATION OF VIGILANCE AND DELAY TASK SCORES

Although the two tasks contained within the GDS tap into different aspects of an individual's functioning, these dimensions are in many ways related. However, there will be many instances, particularly among clinical populations, when scores on the two tasks appear to contradict each other. The purpose of this section is to assist you in formulating interpretations when the two tasks are in accordance as well as when they are in apparent disagreement.

Toward this end, there are some points which must be addressed before interpretive statements are offered. Some of what follows has been covered in the GDS Interpretive Guide. However, a thorough understanding of these issues is essential to valid and effective interpretation.

The Delay Task is, first and foremost, a measure of behavioral suppression and impulse inhibition (delay). It places very few demands upon the subject's alertness or upon his/her ability to sustain attention. A very important aspect of the Delay Task is that it provides information or feedback to the subject. The flashing light and incrementing counter serve the purpose of providing the subject with a "feedback loop". For optimal performance. the subject must incorporate this feedback and use it to monitor and guide responding. In order to do this, most non-impulsive subjects develop cognitive "mediational strategies", usually involving some covert, verbally-mediated behaviors such as counting. Others develop behaviorally-mediated strategies, usually involving some regular pattern of motoric activity. We have observed that these children tend to make less efficient adaptations to this task than do those who employ more sophisticated, covert strategies.

It is important to note that these strategies are self-generated. The Delay Task is essentially subject-paced, in that the apparatus does nothing but record the subject's responding. The subject must pace his or her responding via the employment of these strategies. There is no aspect of the task which requires the subject to "gear up" to respond; therefore, any impulsivity which emerges does so in the absence of external pressure to respond in a rapid or hasty manner.

By contrast, the Vigilance Task requires a relatively high level of arousal on the part of the subject. The subject must achieve and maintain a stance of alertness and vigilance to perform well on this task. This demand is placed upon the subject by the apparatus.

As opposed to the self-paced nature of the Delay Task, the Vigilance Task is instrument-paced. The stimuli rapidly flashing on the screen require that the subject "gear up" for the possibility of responding with the presentation of each new stimulus. In addition, subject performance is in direct response to the instrument. There are no strategies needed to control and monitor responding; the subject simply must be prepared to respond to the correct stimuli at all times during the task.

This level of responding must be maintained in the absence of feedback. The Vigilance Task provides no information whatsoever on the appropriateness of the subject's responding. There is thus no ongoing process of cognitive mediation, as there is in the Delay Task. The subject must simply maintain a level of arousal or vigilance in order to maintain a state of "preparedness to respond". Obviously, optimal performance requires that the subject also suppress tendencies to respond impulsively while in this state of arousal.

Interpretive Statements

NORMAL Delay Task, NORMAL CORRECT (CR) & ERRORS OF COMMISSION (EC) on the Vigilance Task:

	Delay	Vigilance	
	ER	CR	EC
NORMAL	×	×	×
ABNORMAL			

This subject demonstrates an ability to delay and inhibit behavioral responding which is commensurate with his/her peers. This ability to delay is not only adequate in situations when behavioral suppression is the primary requirement, but also when high levels of arousal must be maintained. S/he can perform effectively when alertness and arousal are required, with no evidence of increased impulsivity under these conditions. If this child was referred for ADD, a more extensive psychological evaluation should be considered because problem behaviors do not appear to stem from fundamental deficits in impulse inhibition or sustained attention. Anxiety, emotional difficulties, and/or environmental stress may account for the presenting problems.

NORMAL Delay Task, NORMAL CR & ABNORMAL EC on Vigilance Task:

	Delay ER	Vigilance CR	Vigilance EC
NORMAL	X	X	
ABNORMAL			X

This subject demonstrates an ability to inhibit impulses which is commensurate with his/her peers in situations where behavioral suppression is required for successful performance. S/he can pace and monitor responding under conditions of low arousal and in the presence of informational feedback. The ability to sustain alertness and vigilance is also commensurate with peers. However, under conditions of high motivation & arousal where there is no informational feedback, responding becomes impulsive and poorly controlled. Programming should involve avoiding competitive, high-arousal activities in the classroom. Accuracy should be emphasized over speed of responding, and reinforcement should be used judiciously, as it may heighten arousal levels.

NORMAL Delay Task, ABNORMAL CR & EC on Vigilance Task:

	Delay ER	Vigilance CR	Vigilance EC
NORMAL	X		
ABNORMAL		X	X

This subject demonstrates an ability to inhibit or suppress behavioral responding in self-paced activities where feedback is provided. However, under conditions of high arousal and no feedback s/he demonstrates lapses in sustained attention or alertness in addition to abnormally high levels of impulsivity. In these situations s/he is prone to miss salient aspects of a situation and to respond in a poorly-controlled manner. S/he should be encouraged to utilize strengths (feedback utilization, self-pacing, delay), and interventions should be geared toward helping to incorporate these strengths in situations where vigilance is required. Competitive, speed-oriented exercises should be avoided. S/he should perhaps be allowed extra time to complete tasks such as exams and classroom quizzes.

NORMAL Delay Task, ABNORMAL CR & NORMAL EC on Vigilance Task:

	Delay ER	Vigilance CR	Vigilance EC
NORMAL	X		X
ABNORMAL		X	

This subject demonstrates normal capacity to delay and inhibit responding in situations where responding is self-paced and feedback is provided as well as in situations where high levels of arousal are required and no feedback is provided. However, s/he demonstrates lapses in alertness and sustained attention. S/he should be encouraged to utilize strengths (feedback utilization, self-pacing, delay), and interventions should be geared toward helping them to incorporate these strengths in situations where vigilance is required. S/he should perhaps be allowed extra time to complete tasks such as exams and classroom quizzes.

ABNORMAL Delay Task, NORMAL CR & EC on Vigilance Task:

	Delay ER	Vigilance CR	Vigilance EC
NORMAL		X	X
ABNORMAL	X		

This subject demonstrates a marked tendency to respond impulsively when s/he must pace and monitor responding. Feedback tends to be underutilized and mediational strategies tend to be ineffective. However, under conditions of increased structure (non-subject paced), s/he demonstrates adequate capacity for delay and normal attention and alertness. This child would likely benefit from self-control training programs aimed at teaching cognitively-oriented strategies for delay.

ABNORMAL Delay Task, NORMAL CR & ABNORMAL EC on Vigilance Task:

	Delay	Vigilance	
	ER	CR	EC
NORMAL		X	
ABNORMAL	X		X

This subject demonstrates significant impulsivity, which emerges across many situations. Although vigilance and alertness is commensurate with that of peers, impulsivity is pervasive and may profoundly affect efficiency and effectiveness. S/he will require greatly increased external structure in the classroom, particularly when confronted with tasks which involve self-control and the capacity to inhibit responding. This child would benefit from special programs designed to teach strategies for developing cognitive self-control over behavior. This child will require help in learning how to "stop, look, listen, and think" before acting. Treatment with stimulant medication may be a viable option to consider as part of a general program of educational and psychological management.

ABNORMAL Delay Task, ABNORMAL CR & NORMAL EC on Vigilance Task:

	Delay	Vigilance	
	ER	CR	EC
NORMAL			X
ABNORMAL	X	X	

This subject demonstrates a degree of impulsivity considerably greater than that of his/her peers. This tendency toward impulsive responding is most prominent in situations in which behavioral suppression is required and s/he must pace and monitor his/her responding. This tendency diminishes under conditions of increased structure (when tasks are not self-paced). However, under these conditions s/he exhibits marked lapses in alertness and sustained attention. S/he will thus be prone to miss salient aspects of stimulus situations. Motivation to perform under conditions of high arousal is suspect, as subject may be inhibiting unduly to avoid a torrent of impulsive responding.

ABNORMAL Delay Task, ABNORMAL CR & EC on Vigilance Task:

	Delay	Vigilance	
	ER	CR	EC
NORMAL			
ABNORMAL	X	X	X

This subject demonstrates pervasive impulsivity coupled with deficits in sustained attention. S/he responds in an abnormally impulsive manner across situations, and is prone to miss the salient aspects of a stimulus situation. This child should be considered for treatment with stimulant medication. S/he will require greatly increased external structure in the classroom, particularly when confronted with tasks which involve self-control and the capacity to inhibit responding. Programming should avoid high-arousal situations and include training in cognitive self-control strategies.

The interpretations offered above can be embellished and refined considerably by taking into account the variety o factors within each task as well as by incorporating findings from other sources. Moreover, once the practitioner becomes thoroughly familiar with the nuances of GDS performance and scores, interpretations will take on a much more individualized character.

STANDARD INSTRUCTIONS

Delay Task

You're going to play a game in which you will get a chance to win a lot of points. Do you see this light (pointing to the small red light)? Every time you make this light go on you'll earn a point and this counter (pointing) will keep track of how many points you've won. At the end of the game we'll see how many points you've earned. Now, to make the light go on all you have to do is push this blue button (pointing), and wait a little while before pressing it again. You just press this blue button, wait a while, then press it again. If you press it again too soon, though, you will not get a point, the light won't go on, and you'll have to wait a while before you can press it to get another point. But if you push the button, wait a while, then push it again, you'll get a point every time.

Vigilance Task

Following the Delay Task

Now you're going to play another game. In this game you will see numbers flash on the screen, and I want you to press the blue button every time you see a "9" that comes right after a "1". If the "9" comes after any other number, don't press the button. The only time you should press it is if you see a "9" that comes right after a "1". Now the red light won't go on at all but at the end of the game I'll tell you how many points you won. You will know the game is over when this green light goes on. Do you understand? (Have subject repeat.)

Not Following the Delay Task

This is a game in which you will see numbers flash on the screen quickly. What you need to do in order to win points is to press the blue button every time you see a "9" that comes right after a "1". If the "9" comes after any other number, don't press the button. The only time you should press it is if you see a "9" that comes right after a "1". You will know the game is over when this green light goes on. Do you understand? (Have subject repeat.)

GORDON DIAGNOSTIC SYSTEM

THRESHOLD TABLES

ALL AGES

	NORMAL	BORDERLINE	ABNORMAL
Efficiency Ratio Slope	0 - ±.15		< -.15 or >.15

3-5 YEAR OLD CHILDREN

	NORMAL	BORDERLINE	ABNORMAL
DELAY TASK (N=150)			
Efficiency Ratio			
Total	.61 - 1.00	.43 - .60	0 - .42
Block Variab.	.00 - 0.14	.15 - .25	≥ .26
Single Block	.61 - 1.00		0 - .40
Responses - Total	21 - 66	67 - 89	≤20 or ≥90
Correct - Total	≥ 26	16 - 25	0 - 15
VIGILANCE TASK (N=132)			
Correct			
Total	21 - 29	14 - 20	0 - 13
Single Block	NA		NA
Commissions			
Total	0 - 6	7 - 25	≥ 26
Single Block	0 - 1		≥ 11
Block Variab.	0 - 2.07	2.08-4.99	≥ 5

6-7 YEAR OLD CHILDREN

	NORMAL	BORDERLINE	ABNORMAL
DELAY TASK (N=142)			
Efficiency Ratio			
Total	.72 - 1.00	.42 - .71	0 - .41
Block Variab.	.00 - 0.11	.12 - .20	≥ .21
Single Block	.71 - 1.00		0 - .39
Responses - Total	21 - 63	64 - 75	≤20 or ≥76
Correct - Total	≥ 35	15 - 34	0 - 14
VIGILANCE TASK (N=100)			
Correct			
Total	33 - 45	22 - 32	0 - 21
Single Block	12 - 15		0 - 6
Commissions			
Total	0 - 10	11 - 23	≥ 24
Single Block	0 - 3		≥ 9
Block Variab.	0 - 2.07	2.08-4.99	≥ 5

8-9 YEAR OLD CHILDREN

	NORMAL	BORDERLINE	ABNORMAL
DELAY TASK (N=137)			
Efficiency Ratio			
Total	.78 - 1.00	.52 - .77	0 - .51
Block Variab.	.00 - 0.10	.11 - .23	≥ .24
Single Block	.77 - 1.00		0 - .49
Responses - Total	21 - 65	66 - 75	≤20 or ≥76
Correct - Total	≥ 41	22 - 40	0 - 21
VIGILANCE TASK (N=96)			
Correct			
Total	40 - 45	33 - 39	0 - 32
Single Block	14 - 15		0 - 11
Commissions			
Total	0 - 8	9 - 21	≥ 22
Single Block	0 - 2		≥ 9
Block Variab.	0 - 1.52	1.53-3.78	≥ 3.79

10-11 YEAR OLD CHILDREN

	NORMAL	BORDERLINE	ABNORMAL
DELAY TASK (N=131)			
Efficiency Ratio			
Total	.81 - 1.00	.55 - .80	0 - .54
Block Variab.	.00 - 0.11	.12 - .25	≥ .26
Single Block	.80 - 1.00		0 - .47
Responses - Total	21 - 67	68 - 76	≤20 or ≥77
Correct - Total	≥ 45	32 - 44	0 - 31
VIGILANCE TASK (N=88)			
Correct			
Total	42 - 45	39 - 41	0 - 38
Single Block	14 - 15		0 - 12
Commissions			
Total	0 - 3	4 - 15	≥ 16
Single Block	0 - 1		≥ 5
Block Variab.	0 - 1.15	1.16-1.99	≥ 2

12-16 YEAR OLD CHILDREN

	NORMAL	BORDERLINE	ABNORMAL
DELAY TASK (N=156)			
Efficiency Ratio			
Total	.86 - 1.00	.71 - .85	0 - .70
Block Variab.	.00 - 0.10	.11 - .20	≥ .21
Single Block	.86 - 1.00		0 - .66
Responses - Total	21 - 66	67 - 75	≤20 or ≥76
Correct - Total	≥ 47	39 - 46	0 - 38
VIGILANCE TASK (N=156)			
Correct			
Total	43 - 45	40 - 42	0 - 39
Single Block	15		0 - 12
Commissions			
Total	0 - 3	4 - 11	≥ 12
Single Block	0		≥ 5
Block Variab.	0 - .99	1.00-1.9[illegible]	≥ 2

PRESCHOOL INSTRUCTIONS DELAY TASK

"You are going to play a game in which you will get a chance to win a lot of points, not just one or two, but a whole bunch of points. Do you see this light (pointing to the small red light)? Every time you make this light go on, you'll earn a point, and this counter (pointing) will keep track of how many points you've won. At the end of the game, we'll see how many points you've earned. Now, to make the light go on, all you have to do is push this blue button (pointing), and wait a little while, then press it again. If you press it again too soon, though, you will not get a point, the light won't go on, and you'll have to wait a while before you can press it to get another point. But, if you push the button, wait a while, then push it again, you'll get a point every time.

Do you have any questions?

I'll sit back here and wait until you've finished. You will know the game is over when the green light comes on and a buzzer goes off. If you have any questions or want to talk about the game, I want you to wait until the game is over, and we can discuss it then.

Are you ready (said with enthusiasm)? OK, let's begin (or "ready, set, go!").".

VIGILANCE TASK

Instruction And Practice Session

"Now to play the next game, you need to know what the number "1" looks like. Can you show me a "1" (showing them the flash card with LED-style numbers)? Can you show me a "7"? Can you show me a "1" again? (The child is asked to point to the number "7" because it looks similar to a "1", so that it is necessary to make sure that the child can differentiate between the two numbers. Since some children are unable to discriminate between the numbers, you should try to teach them. If the child still cannot pick out the "1", then the task cannot be administered.)

When you play this game, you are going to see numbers flash on the screen, and I want you to press the blue button only when you see a "1". Only press the blue button if you see a "1" and not when any other number flashes. Are you ready? Begin."

Assessment Session

"Now you are going to play this game for a longer time. I want you to press the blue button only when you see a "1". Only press the button if you see a "1". I'll sit here and wait until you're done. You'll know when the game is over, because the green light will come on, and you will hear a beep. If you have any questions, or want to talk about the game, I want you to wait until after the game is over, and we can discuss it then."

DELAY TASK SETTINGS

- *Set TASK SELECTOR SWITCH to* ***DELAY TASK***
- ***INTERVAL*** *Thumbwheel - 4 seconds*
- ***BLOCK LENGTH*** *Thumbwheel - 90 seconds*
- *Press four Block* ***PROGRAMMER*** *buttons*

ONE MODE/VIGILANCE TASK SETTINGS

Instruction And Practice Session

- *Set TASK SELECTOR SWITCH to* ***VIGILANCE TASK***
- *Set* ***Mode Selector*** *to "1"*
- ***INTERVAL*** *Thumbwheel - 2 seconds*
- ***BLOCK LENGTH*** *- 20 seconds*
- *Press* ***BLOCK 1*** *Programmer button*
  - a. ***BLOCK LENGTH*** *Thumbwheel - 0 seconds*
  - b. ***INTERVAL*** *Thumbwheel - 2 seconds*
  - c. *Press* ***BLOCK 2*** *and* ***3*** *Block Programmer buttons*

Assessment Session

- *Press* ***RESET*** *button*
- ***BLOCK LENGTH*** *Thumbwheel - 120 seconds*
- ***INTERVAL*** *Thumbwheel - 2 seconds*
- *Press first 3* ***BLOCK*** *Programmer buttons*

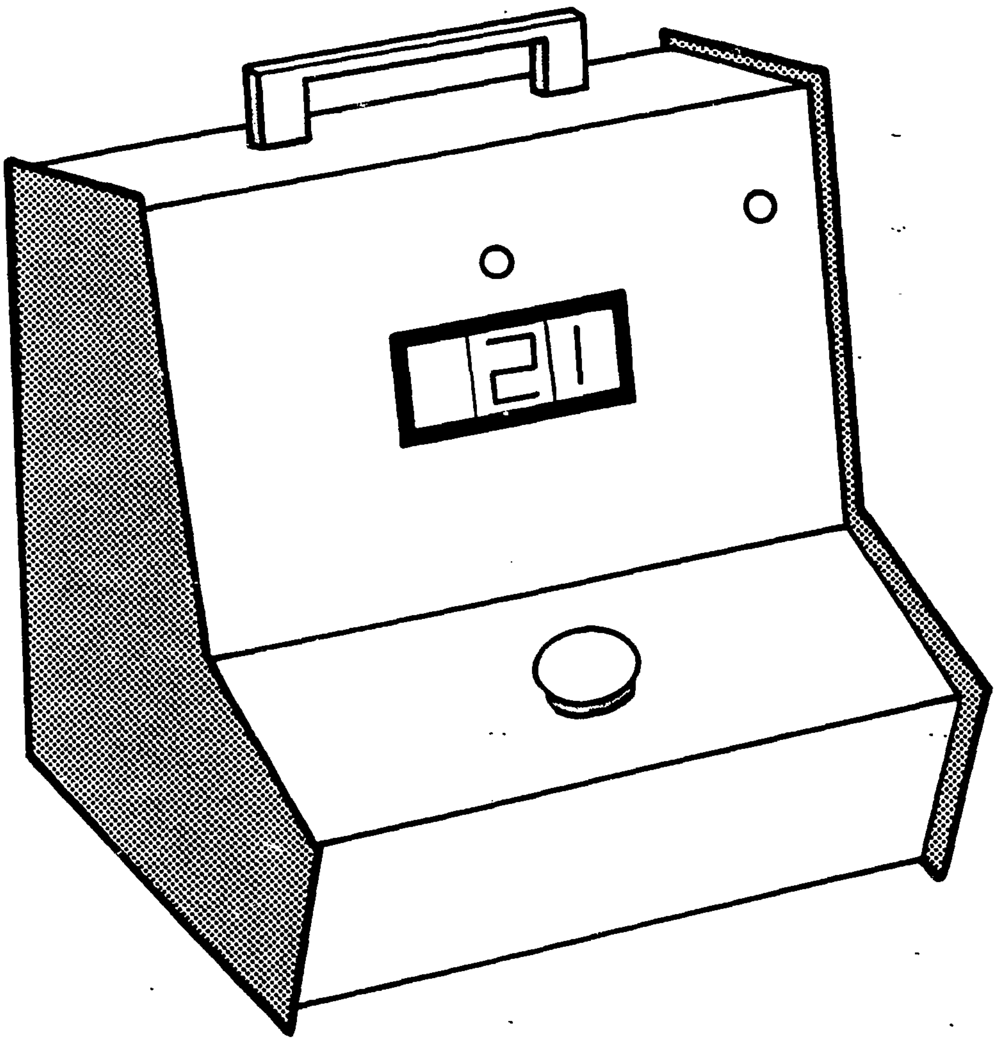
REFER TO MANUAL FOR DETAILED INSTRUCTIONS

GORDON DIAGNOSTIC SYSTEM
PRESCHOOL CHILDREN
THRESHOLD TABLES

Efficiency Ratio Slope	NORMAL 0 - $\pm$.15	BORDERLINE	ABNORMAL < -.15 or >.15
3-5 YEAR OLD CHILDREN			
	NORMAL	BORDERLINE	ABNORMAL
DELAY TASK (N=150)			
Efficiency Ratio			
Total	.61 - 1.00	.43 - .60	0 - .42
Block Variab.	.00 - 0.14	.15 - .25	$\geq$ .26
Single Block	.61 - 1.00		0 - .40
Responses - Total	21 - 66	67 - 89	$\leq$20 or $\geq$90
Correct - Total	$\geq$ 26	16 - 25	0 - 15
VIGILANCE TASK (N=132)			
Correct			
Total	21 - 29	14 - 20	0 - 13
Single Block	NA		NA
Commissions			
Total	0 - 6	7 - 25	$\geq$ 26
Single Block	0 - 1		$\geq$ 11
Block Variab.	0 - 2.07	2.08-4.99	$\geq$ 5

THE GORDON DIAGNOSTIC SYSTEM

DATA ANALYSIS PROGRAM



CLINICAL DIAGNOSTICS INC
1536 Cole Blvd., Golden, Colorado 80401
Suite 350 • 800-521-4503

THE GORDON DIAGNOSTIC SYSTEM DATA ANALYSIS PROGRAM

Ernest M. Post, M.D.
Cooper Hospital/University Medical Center, Camden, New Jersey

Michael Gordon, Ph.D.
S.U.N.Y. Upstate Medical Center, Syracuse, New York

Table of Contents

Organization of this Manual

The GDS Data Analysis Program (DAP) was developed to enhance and facilitate the operation of the Gordon Diagnostic System (GDS). This manual explains how to run the DAP on your computer. Since many DAP users have never before touched a computer, we have tried to keep the terminology and instructions as simple as possible.

It is not our intention to teach you to use your personal computer; that is the function of the operating manual provided with your machine. Most of the operation of the DAP is the same regardless of which computer you are using. Those features which vary among machines are mentioned briefly in the text and explained further in the Appendix.

I. **GETTING STARTED** tells you how to connect the GDS to your computer and how to start running the DAP.

II. **CONVERSING WITH THE DAP** describes in general terms how the computer asks for information and how you should respond.

III. **STEP BY STEP** tells you how to direct the DAP to collect, interpret, and save the results generated by the GDS.

IV. **FUNCTIONS MENU** presents the first half of the DAP features you can access using the function keys.

V. **STORED RESULTS** describes the other features of the DAP, which let you recall information stored during previous testing sessions.

APPENDICES

A. **SYSTEM REQUIREMENTS** is a general description of what you need to connect the GDS to a computer.

B. **SPECIFIC COMPUTERS** contains the necessary details for running the DAP on a variety of computers.

C. **ORGANIZATION OF DATA FILES** describes the file structure so that you may access the files from other programs.

D. **MAKING COPIES** is "must reading" for the computer novice. It discusses the need to keep duplicate "copies" of all important documents.

Acknowledgements

We would like to thank everyone who helped refine The Data Analysis Program and clarify this manual. We are particularly grateful for the contributions of Dr. F. Daniel McClure and the thorough evaluations by Ms. Barbara Bass and Ms. Mary Gallaghan. As always, we appreciate the gracious tolerance of our wives and children.

January 26, 1984

I. GETTING STARTED

Before you can start using the DAP, you need to have the appropriate pieces of equipment properly connected. The details are described in Appendix A: **SYSTEM REQUIREMENTS**, but basically you need the Gordon Diagnostic System (GDS), an appropriate computer, and the DAP. Depending on your computer, the DAP may have been provided for you on a "floppy disk". tape cassette, or "magnetic bubble" cartridge.

1. Setting up the Equipment

First, if your computer has separate components, you should connect them following the instructions in your owner's manual. Next, the GDS and the computer need to be connected by a cable (see Appendix A), which should have been purchased with your computer. (If the connectors on the two ends of the cable are the same, it doesn't matter which end goes where.) There is only one possible place to connect the cable to the GDS. Some computers may have more than one place where the other end of the cable could be connected; be sure you connect the cable to the part of the computer called the "serial port", the "asynchronous port", or the "RS-232 connector". (The three terms are essentially synonymous.)

2. Getting the DAP into your computer

Once you have everything connected, you need to "load" the DAP into your computer. If it is stored on a floppy disk, you need to insert the disk in your computer as described in the owner's manual. If the DAP is in a magnetic bubble cartridge. the cartridge needs to be in place in the computer. Unlike the floppy disk, it can be left in place when the machine is turned off, so it need not be inserted for each session. If the DAP is stored on a tape cartridge, it needs to be "loaded" from the cartridge to the computer before it can run. For some computers, once loaded, the DAP will stay in the computer and need not be loaded for each session.

3. Starting the DAP

As with the previous section. there are differences among computers as to the procedures for starting the DAP. In some computers, the DAP is set up to start running automatically. In others, you will need to tell the computer that you want to run the DAP, as opposed to some other program you may have. Specific details of how to start the DAP in each machine have again been left for Appendix B.

II. CONVERSING WITH THE DAP

The DAP lets you decide what to do by displaying a series of "prompts" or questions, to which you must respond by pressing one or more keys on the computer keyboard. The DAP only recognizes capital letters; if your computer has a lower case mode, you should press the "CAPS", "CAPS LOCK", or other appropriate key to make all the letters you type appear in UPPER CASE. (If you're unsure, see Appendix B.)

There are several types of messages (prompts) displayed by the DAP to request a response from you. Each is described below and followed by an example:

1. Yes/No Choices

These are generally self-explanatory. For example, the very first prompt to appear is:

PRINTER READY (Y/N)?>

You are expected to respond by pressing either the "Y" or "N" key on the computer keyboard. For this and several other responses only a single key needs to be pressed. To repeat, RESPONSES MUST BE IN CAPITAL LETTERS OR THE DAP WON'T ACKNOWLEDGE THEM.

2. Other two-option choices

These prompts are similar to the (Y/N) prompts. One prompt using this format is:

QUICK OR STANDARD LIST (Q/S)?>

You would respond with the appropriate letter, Q or S, as will be explained later.

3. Function Keys

If your computer (see Appendix B) has keys labelled F1, F2, F3, etc., these are called function keys because they are used to select the function which you wish the program to perform. The functions of the DAP will be explained in detail in the next section, but are listed below:

Key Function
F1 IDENTIFY NEW SUBJECT
F1 ENTER RESULTS
F3 REDISPLAY CURRENT RESULTS
F4 LIST STORED SUBJECTS
F5 RECALL STORED RESULTS
F6 DELETE STORED RESULTS
F7 END SESSION

There is one other function which doesn't appear on this list, but which can be implemented whenever this list is displayed by the computer (except on the Apple). As will be stated by the computer, you can begin running one of the GDS tasks; when the computer displays the list of functions it is also ready to start receiving information about the tasks from the GDS.

4. Print Option

The computer displays the results of the GDS testing and their interpretations on its screen in several segments as described in the next chapter. At the end of each of these segments the following prompt appears:

PRINT/CONTINUE/SKIP RESULTS (P/C/S)?

If you respond by pressing "P", and you previously told the DAP you had a printer ready, the computer will print exactly what just appeared on its screen to give you a paper copy for future reference. If you press "C", the computer will go on to the next step, which is often displaying more results. If you press "S", the computer will bypass displaying the remainder of the results and give you a choice of the "functions" again. In some instances, the pressing either "C" or "S" will have the same effect. You never need to worry that pressing "S" will result in losing data; it skips the display of the results, but does not affect their storage.

5. Requests for specific information

Sometimes the computer needs more than a single letter response from you. For example, when the computer wants to know the name of the examiner, the following prompt appears:

EXAMINER NAME>

You then type the examiner's name exactly as you want it to appear on everything that the computer will print about that particular subject. You may use last name first, abbreviations, or nicknames if you so choose. Whatever you enter is exactly what the computer will give back to you. For the computer to know that you have finished typing the name. YOU MUST PRESS the <ENTER> or <RETURN> key (sometimes marked by an arrow pointing down, then left). If you make a mistake while typing your response and you haven't yet hit the <ENTER> key, typing errors can be corrected by pressing the "Backspace" button. Its position and operation vary among computers; consult your computer manual to learn how to use it. Even if you have pressed <ENTER>, you will still have an opportunity to make changes later on.

6. Subject Name

There are several prompts similar to the one just described, but one is slightly different and is described below:

Request for subject name – This prompt can appear in two forms. When it appears as

SUBJECT NAME>

it is the same as other requests for information. What you get back will be exactly what you put into the computer. When it appears as:

ENTER SUBJECT NAME>

it is more flexible. This form is used when you are asking the computer to do something with data that has already been stored. You do not need to enter the subject's full name; any part of it will do. For example, if the subject is MIKE SMITH, you only need to answer this prompt with MIKE. The computer will search for the record of anyone named MIKE and ask you if it found the right Mike before it goes on. However, the computer cannot find the record if you originally entered the name as MICHAEL SMITH. If you forget how you entered it, you could ask for MI, but be careful because the computer will also find MIRIAM and TAMMI that way. If you ask for SMITH, it will try all subjects named SMITH until it finds the one you want.

7. Escape Key

Because it takes the computer several seconds to generate some of the displays described below, the "long" ones are provided with an "escape" mechanism. If you press the <ESCAPE> key **(See Appendix B)** the routine will be interrupted and the computer will go on to the next step. If you "escape" before or during the screen display, you will still have an opportunity to print the display on your printer. The displays which use this feature are the graphics, the Interpretations, and the Task Monitoring Data.

III. STEP BY STEP

This chapter gives detailed instructions for using the DAP. You may want to try each step as you read about it. The sections of this chapter are as follows:

1. Printer Ready Prompt
2. Starting a Task
3. Subject Identification
4. Data Display

1. Printer Ready Prompt

Once you have started the DAP, as described in GETTING STARTED, you should hear a beep and see:

GDS Data Analysis Program

(C) 1984 Clinical Diagnostics, Inc.
Serial #840120

USE CAPITAL LETTERS ONLY
PRINTER READY (Y/N)?>

Answer this question carefully. If the printer is connected and ready, type Y. If you type Y and the printer is not ready, the program may not be able to proceed. The following message will appear:

IF THIS STAYS ON PRINTER'S NOT READY.

PRESS <BREAK>THEN <F8>

Pressing <BREAK> can require that you press two keys rather than one for some computers **(See Appendix B).**

If you type N in response to the PRINTER READY prompt, the DAP will not print anything on the printer until you have started again from the beginning. If you don't have a printer ready when you are administering the GDS, don't worry. The data collected from the GDS can be saved and printed later if desired.

2. Starting a Task

Once you have responded to the PRINTER READY prompt, the following will appear on your computer:

START TASK OR PRESS A FUNCTION KEY
F1 IDENTIFY NEW SUBJECT
F2 ENTER RESULTS
F3 REDISPLAY CURRENT RESULTS
F4 LIST STORED SUBJECTS
F5 RECALL STORED RESULTS
F6 DELETE STORED RESULTS
F7 END SESSION

The computer is now ready to accept data from the GDS. You do not have to do anything else with the computer itself. Simply begin the GDS task, according to procedures outlined in the GDS Instruction Manual. (With the Apple computer, you will need to press F2 and then E before the computer is ready to accept data from the GDS.)

While the task is being performed, Task Monitoring Data will appear on the computer screen. You may observe the progress of the task, but it may be more valuable to keep your attention on the child's behavior at this time.

The data that is displayed is different for the two tasks. For the Delay Task the following is displayed each time the blue button is pressed:

INTERVAL #n WAS x.x SEC

This tells you how long it was since the previous button press and gives a picture of the child's progress. For the Vigilance Task, the display is a series of digits separated by a blank space or by a hyphen (dash). The digits are those displayed on the GDS for the task. The presence of a dash indicates that the blue button was pressed after that digit. If there is no dash, the button was not pressed.

While observing these results can be informative, there is no need to take notes because a summary of them is displayed by the computer at the conclusion of the task. (See Data Display below.)

3. Subject Identification

When the task has been completed, i.e., when the GDS beeps and the green Game Over light comes on, the following prompt will appear:

nn OF mm POSSIBLE SUBJECTS STORED

SUBJECT NAME>

The first line tells you how much storage space you have left. If your space is full, you will be so informed and offered the opportunity to delete a subject **(See Chapter V. F6 DELETE STORED RESULTS).**

In response to the SUBJECT NAME prompt you should enter the name of the subject exactly as you want it to appear when the computer displays or prints the task results. You may use last name first, with or without a comma. You are limited to a maximum of 25 letters, spaces, numbers or punctuation marks (collectively termed characters). After you type the name, press the <ENTER> key to tell the computer you have finished entering the name. The computer will then [illegible]:

SEX (M/F)>

You should type M or F. and then press <ENTER>. The next prompt is:

BIRTH (MM/DD/YY)>

to which the appropriate response would be 05/09/76 for a child born on May 9, 1976. The computer does not check that you have entered the date properly, but the zeros before the 5 and 9 are needed for the computer to determine the child's age. Next is:

EXAMINER NAME>

and again you enter the name as you wish, with a maximum of 13 characters. After that comes:

COMMENT>

Here you can respond with up to 13 characters of information (e.g., you may want to identify the child's school or reason for referral). The last prompt in this series is:

TEST DATE (MM/DD/YY)>

You need to respond in the same format as that used for the birth date. The computer then asks:

CHANGE THESE ENTRIES (Y/N)?>

If you answer Y, then all of the subject identification prompts are repeated and all the information needs to be reentered. If you answer N, your responses are stored and the computer goes on to the next step.

You may leave the sex, examiner and comment entries blank by pressing <ENTER> without typing anything. If you leave either date blank, the computer will be unable to determine the subject's age. If you leave the subject's name blank, the data CANNOT be stored. The results will still be displayed, but cannot be retrieved later on.

There is an earlier opportunity to enter all of the subject identification data, but that won't be discussed until **Chapter IV, Functions Menu.**

9

4. Data Display

After you have administered the task and entered the subject information, the computer automatically starts to display the task results and their interpretation. Samples of each display are shown and described below. The first display repeats the subject identification information to give you an opportunity to print it.

```
                  SUBJECT NUMBER 6

SUBJECT NAME:                          MICHAEL SMITH
SEX (M/F):                                         M
BIRTH (MM/DD/YY):                           05/09/76
EXAMINER NAME:                           JOHN PSYCHE
COMMENT:                                    R/O A.D.D.
TEST DATE (MM/DD/YY):                       01/10/84
PRINT/CONTINUE/SKIP RESULTS (P/C/S)?
```

Choose one of the alternatives, P. C, or S, as described in **Chapter II, Conversing with the DAP.**

The next display shows the Summary Statistics, which includes the Parameters used when administering the task and the scores achieved on the task. These scores are the same as those that can be read from the back panel of the GDS using the Data Selector. The displays from the Delay and Vigilance Tasks are slightly different, so only the Delay Task is illustrated here. An illustration of the Vigilance Task display is included at the end of this chapter.

BLOCK:	1	2	3	4	TOTAL
INTERVAL	6	6	6	6	
LENGTH	120	120	120	120	480
REWARDS	10	8	10	4	32
RESPONSES	18	23	23	27	91
EFFIC RATIO	.56	.35	.43	.15	.35

All of the terms in the above example are described in the Instruction Manual, Interpretive Guide, and Interpretive Supplement. To briefly review, the Interval and Length are the Task Parameters you specified. Rewards is the number of button presses that were made after waiting a sufficient period of time. Responses is the total number of button presses and Efficiency Ratio is determined from dividing the number of Rewards by the number of Responses. These results are displayed for each time Block and for the total Task.

The next display is a graphic presentation of the Summary Statistics.

```
                  REWARDS
     !                                   !
25 _ !                                   !
     !                                   !
20 _ !                                   !
     !                                   !
15 _ !                                   !
     !                                   !
10 _ !                                   !
     !   XXX     XXX     XXX             !
 5 _ !   XXX     XXX     XXX             !
     !   XXX     XXX     XXX     XXX     !
 0 _ !   XXX     XXX     XXX     XXX     !
     -------------------------------------
          1       2       3       4
                  BLOCKS

                 RESPONSES
     !                                   !
50 _ !                                   !
     !                                   !
40 _ !                                   !
     !                                   !
30 _ !                                   !
     !                           XXX     !
20 _ !           XXX     XXX     XXX     !
     !   XXX     XXX     XXX     XXX     !
10 _ !   XXX     XXX     XXX     XXX     !
     !   XXX     XXX     XXX     XXX     !
 0 _ !   XXX     XXX     XXX     XXX     !
     -------------------------------------
          1       2       3       4
                  BLOCKS

                EFFIC RATIO
     !                                   !
 1 _ !                                   !
     !                                   !
.8 _ !                                   !
     !                                   !
.6 _ !                                   !
     !   XXX                             !
.4 _ !   XXX             XXX             !
     !   XXX     XXX     XXX             !
.2 _ !   XXX     XXX     XXX             !
     !   XXX     XXX     XXX     XXX     !
 0 _ !   XXX     XXX     XXX     XXX     !
     -------------------------------------
          1       2       3       4
                  BLOCKS
```

For each graph, the bar extends above the horizontal line at the left if the value exceeds the corresponding number. For example, the Rewards for Block 1 was 10, which is greater than 7.5, but not greater than 10, so that the bar extends up to, but not above the line at 10. The graphics can be bypassed by pressing the <ESCAPE>key.

The next display shows interpretations based upon the subject's scores. (For the Vigilance Task, the interpretations also depend on the subject's age.) The logical basis of the interpretation is described in the Supplement to the Interpretive Guide.

TOTAL SCORES INDICATE:
Significant inability to delay
Shows clear evidence of impulsivity
Interpret protocol with extra caution because of very high responsivity

BLOCK SCORES INDICATE:
Never reaches normal range of behavioral delay
Block variability in borderline range
ay be more inconsistent than peers in level of impulsivity exhibited
Performance deteriorates during session
Will become more impulsive over time

EDUCATIONAL PROGRAMMING:
Teach cognitive mediational strategies
Emphasize correctness over speed
Encourage self-monitoring
Evaluate for possible motivational/emotional factors
Extremely unable to suppress responding
Will need increasing structure as task progresses

The interpretations are only available if the Standard Parameters were used and, for the Vigilance "1/9" task if the child is at least 6 years old. This routine can be bypassed by pressing the <ESCAPE>key.

The final display is a summary of the Task Monitoring Data and gives the details of the subject's performance on the task. For the Delay Task it appears as follows:

INITIAL RESPONSE LATENCY = 01.7 SEC

INT.	BLOCK: 1	2	3	4	TOTAL
0 SEC	0	0	0	0	0
1 SEC	1	0	0	0	1
2 SEC	1	2	1	2	6
3 SEC	4	4	3	6	17
4 SEC	1	5	5	9	20
5 SEC	1	4	4	6	15
6 SEC	4	3	4	2	13
7 SEC	4	3	5	0	12
8 SEC	1	1	1	1	4
9 SEC	0	1	0	0	1
10 SEC	0	0	0	0	0
11 SEC	0	0	0	0	0
12 SEC	0	0	0	1	1
13 SEC	0	0	0	0	0
14 SEC	0	0	0	0	0
15 SEC	0	0	0	0	0
16 SEC	0	0	0	0	0
17 SEC	0	0	0	0	0
18 SEC	0	0	0	0	0
19 SEC	0	0	0	0	0
>19 SEC	0	0	0	0	0

This display tells you the number of times the button was pressed after the stated interval, for each Block and for the total Task. This can be bypassed by pressing the <ESCAPE> key.

After this display, the computer returns to the same display (the Functions Menu) as that shown in **Section 2, Starting a Task** and which is explained further in the next chapter.

Vigilance Task Display Example

SUMMARY STATISTICS

BLOCK:	1	2	3	TOTAL
INTERVAL	1	1	1	
LENGTH	180	180	180	540
CORRECT	12	7	11	30
OMISSIONS	3	8	4	15
COMMISSIONS	7	18	21	46

The terms in the above example that differ from the Delay Task are: Correct, Omissions, and Commissions. As explained in the Interpretive Guide, Correct is the number of times the button was pressed after the 1/9 sequence appeared. Omissions is the number of times the subject erred by missing the 1/9 sequence. Commissions is the number of errors the subject made by pressing the button when the 1/9 sequence had not appeared.

GRAPHICS

The displays for the Vigilance Task are similar to those generated for the Display Task.

```
                        CORRECT
     !                                          !
15 _ !                                          !
     !                                          !
12 _ !                                          !
     !   XXX                         XXX        !
 9 _ !   XXX                         XXX        !
     !   XXX                         XXX        !
 6 _ !   XXX          XXX            XXX        !
     !   XXX          XXX            XXX        !
 3 _ !   XXX          XXX            XXX        !
     !   XXX          XXX            XXX        !
 0 _ !   XXX          XXX            XXX        !
     ------------------------------------------------
          1            2              3
                     BLOCKS
```

```
                        OMISSIONS
     !                                          !
15 _ !                                          !
     !                                          !
12 _ !                                          !
     !                                          !
 9 _ !                                          !
     !                XXX                       !
 6 _ !                XXX                       !
     !                XXX                       !
 3 _ !                XXX            XXX        !
     !   XXX          XXX            XXX        !
 0 _ !   XXX          XXX            XXX        !
     ------------------------------------------------
          1            2              3
                     BLOCKS
```

```
                        COMMISSIONS
     !                                          !
30 _ !                                          !
     !                                          !
24 _ !                                          !
     !                                          !
18 _ !                               XXX        !
     !                XXX            XXX        !
12 _ !                XXX            XXX        !
     !                XXX            XXX        !
 6 _ !   XXX          XXX            XXX        !
     !   XXX          XXX            XXX        !
 0 _ !   XXX          XXX            XXX        !
     ------------------------------------------------
          1            2              3
                     BLOCKS
```

INTERPRETATIONS

SUBJECT'S AGE 7.5 YEARS

TOTAL SCORES INDICATE:
Normally alert and vigilant
No evidence of inattentiveness on this task
Check Delay Task performance for level of impulsivity in self-paced condition
Normal Delay Task performance would indicate ability to integrate feedback to stabilize responding
Ability to suppress responding under conditions requiring sustained attention is in borderline range

BLOCK SCORES INDICATE:
Performance deteriorates during session
Behavioral suppression never reaches normal range

EDUCATIONAL PROGRAMMING
Full psychological evaluation may be indicated

TASK MONITORING DATA

SEQ.	BLOCK: 1	2	3	TOTAL
X19	12	7	11	30
XX1	0	1	2	3
19X	2	6	7	15
XX9	1	0	0	1
X9X	0	1	1	2
X1X	1	2	2	5
XXX	3	8	9	20

This display shows you the sequence of digits that preceded each of the button presses. An X in this table indicates that any digit may have appeared. For example a button press registered in the X19 row is one which correctly followed the presentation of the "1" and "9" digits in sequence. A score in the "19X" row means the three digits preceding the button press could have been 1, 9, and 3 or 1, 9, and 7, etc. Instead of pressing the button right after the "1/9" presentation, the subject pressed the button after the next digit was displayed. The XXX refers to button presses that were unrelated to either a 9 or a 1. In the first block of the above example, the subject pressed the button correctly after 12 "1/9" presentations (X19), on two occasions one digit too late (19X), once when alerted by a 1 followed by a digit other than 9 (X1X) and three times after a digit that was unrelated to the appearance of a 1 or 9.

IV. FUNCTIONS MENU

The display you see after the PRINTER READY prompt, after the displays, or after the completion of any "function" is as follows:

START TASK OR PRESS A FUNCTION KEY
F1 IDENTIFY NEW SUBJECT
F2 ENTER RESULTS
F3 REDISPLAY CURRENT RESULTS
F4 LIST STORED SUBJECTS
F5 RECALL STORED RESULTS
F6 DELETE STORED RESULTS
F7 END SESSION

To perform any of these functions you simply press the appropriate function key, usually marked on the keyboard as F1, F2, etc. On those computers which do not have function keys, such as the Apple, you press the appropriate number key instead. For example if you wanted to identify a new subject, you would press the 1 key when the function name appeared. Functions F1-F3 and F7 will be described in this chapter. Those dealing with stored data, F4-F6, will be discussed in the next chapter.

F1 IDENTIFY NEW SUBJECT

The purpose of this function is to allow you to identify the subject before the task has been performed rather than afterwards. The choice is yours and does not affect how the data is displayed, stored or retrieved. The prompts and responses for this function have already been explained in **Chapter III, Section 3, Subject Identification.**

F2 ENTER RESULTS

For some computers, including the Apple, this function is necessary to let the GDS transfer data to the computer electronically. For all computers, this function gives you access to a new capability not previously described. It lets you manually enter the results of a GDS task into the computer, which can then interpret and store your data. You can use some of the DAP features even if you do not have your computer connected during the administration of the GDS.

When you press F2, the following prompts appear:

ELECTRONIC OR MANUAL DATA ENTRY (E/M)?>

If you respond E, the computer is ready and waiting for data to come from the GDS. (Most computers, except the Apple, are also ready to receive data whenever you see the functions menu.) If you change your mind about running a task, pressing the RESET button ON THE GDS will return you to the function menu.

If you respond M and have not just identified a subject, the computer will show:

YOU MUST IDENTIFY A SUBJECT FIRST

SUBJECT NAME>

and proceed through the usual sequence of prompts to identify the subject. If you press <ENTER> without typing a subject name, the computer will proceed through the next set of prompts, but will not store your data.

Once the subject has been identified, the computer will ask:

DELAY OR VIGILANCE TASK (D/V)?>

to which you respond with D or V depending on the data you are entering.

If you previously stored a task for a subject and try to do so again, the computer will say:

SUBJECT ALREADY HAS THAT TASK STORED

You will need to re-enter the subject identification so the computer treats it as a new subject.

It then asks:

SAME PARAMETERS FOR ALL BLOCKS (Y/N)?>

If you ran the task with the same Interval and Block length for each block then answer Y. The computer will respond:

ENTER INTERVAL FOR BLOCK 1>

For the Delay Task, the standard interval is 6 (sec); for the Vigilance Task it is 1 (sec). The next prompt is:

ENTER LENGTH OF BLOCK 1>

The standard lengths for the Delay and Vigilance Tasks are 120 and 180 (sec), respectively. If you did not use the same parameters for all blocks, the computer will prompt you to enter the Intervals and Lengths for Blocks 2-4. Otherwise it assumes that the Block 1 Interval and Length you entered apply to the other blocks. (If you press <ENTER> after either the INTERVAL or LENGTH prompts, you will be returned to the functions menu without completing this routine.)

The computer will then display:

BLOCK:	1	2	3	4	TOTAL
INTERVAL	6	6	6	6	
LENGTH	120	120	120	120	480

ARE THESE CORRECT (Y/N)?>

If you respond N it will start over at the INTERVAL prompt.

Otherwise it will go on to ask:

DATA FOR SELECTOR POSITION 1>

Your response should be the number that you read from the back of the GDS when the Data Selector was at 1. The DAP will not let you press <ENTER> without typing a number and will not accept any entries greater than 99. After you respond, the computer will prompt you to enter the data for selector positions 2-10. When you finish that the computer will display the 10 numbers you entered plus appropriate sums or quotients, such as:

BLOCK:	1	2	3	4	TOTAL
REWARDS	11	12	10	13	46
RESPONSES	13	15	12	16	56
EFFIC RATIO	.85	.80	.83	.81	.82

ARE THESE CORRECT (Y/N)?>

Again you have the option of reentering the 10 numbers if you made a mistake. If you made no error, pressing Y starts the computer on the display routines to let you print the data and give you interpretations.

F3 REDISPLAY CURRENT RESULTS

The purpose of this function is to let you view the results again and to give you another opportunity to print them if you need to. The displays produced are exactly the same as those described in **Chapter III, Section 4, Data Display.**

F7 END SESSION

When you have finished using the computer, you should press F7. Failure to do so may result in the computer's not storing the last information it handled. If you press F7 and then decide you need to run the DAP again, pressing F7 will restart the DAP. (To restart the DAP on computers which do not have function keys, such as the Apple, you need to type RUN and press the RETURN key instead.)

V. STORED RESULTS

One of the valuable features of the DAP is that it stores the results obtained from administering the GDS. This enables you to print the results at a time and place separate from the task administration. It also allows you to collect sets of results to investigate one subject's responses over time, or differences among diagnostic groups. If you choose to use this feature, you will need to learn some additional steps to retrieve your data.

The data is stored in a floppy disk, magnetic bubble cartridge, or in computer memory, depending on the computer you are using. However, the operation of the functions has been designed to be the same regardless of the storage medium. The only difference is that the capacity to store results varies with the medium as is described in **Appendix B.**

The functions needed to manipulate stored results are F4-F6 on the function menu. The operation of each of them is described below. (Functions F1-F3 and F7 were described in the previous chapter.)

START TASK OR PRESS A FUNCTION KEY
F1 IDENTIFY NEW SUBJECT
F2 ENTER RESULTS
F3 REDISPLAY CURRENT RESULTS
F4 LIST STORED SUBJECTS
F5 RECALL STORED RESULTS
F6 DELETE STORED RESULTS
F7 END SESSION

F4 LIST STORED SUBJECTS

This function produces a list of the names of the subjects that you have stored. The names can be listed in either a quick or standard form as determined by your response to:

QUICK OR STANDARD LIST (Q/S)?>

If you respond Q, a typical list would be:

1 MICHAEL SMITH	D	V9
2 TAMMI JONES	D	
3 PUBLIC, JOHN Q.		V1

PRINT/CONTINUE/SKIP RESULTS (P/C/S)?

The "D" and "V" notations tell you which tasks are stored for each subject. For example, Michael has both tasks (Delay and Vigilance "1/9") stored, but Tammi has only the Delay Task.

The last line, as always, gives you the option of printing the list. The numbers are assigned by the computer and are only temporary. If you delete a subject (see below), that subject's number is reassigned to the next one on the list and all the numbers from there down are changed.

If instead of Q you had responded S, you would see:

SUBJECT NUMBER 1

SUBJECT NAME:	MICHAEL SMITH
SEX (M/F):	M
BIRTH (MM/DD/YY):	05/09/76
EXAMINER NAME:	JOHN PSYCHE
COMMENT:	R/O ADD
TEST DATE (MM/DD/YY):	01/10/84

PRINT/CONTINUE/SKIP RESULTS (P/C/S)?

After you respond P or C to this prompt, the same information will be displayed for the next subject. If you answer S then you will be returned to the function menu. You can repeat this listing function as often as you like.

F5 RECALL STORED RESULTS

This is the function that allows you to bring back into the computer results which it filed away earlier. To do so you will have to tell the computer whose results you want and whether you want the Delay or Vigilance Task. The prompts are as follows:

ENTER SUBJECT NAME>

You should read **Chapter II, Section 6, Subject Name** to learn the optional ways you can answer this prompt. The computer will find the first subject whose name matches your response to the above prompt. It will display that subject's identifying information, as shown in the example under standard listing, except that the last line will be:

IS THIS THE CORRECT SUBJECT (Y/N)?>

If it is not the correct one and you answer N, the computer will find the next subject in the list whose name matches what you requested. If the computer does not find the match that you want, it will say so. If you have difficulty, use F4 LIST STORED SUBJECTS to help you find how you entered the name of the subject you are looking for.

If you answer Y, the computer will ask:

DELAY OR VIGILANCE TASK (D/V)?>

Answer D or V as needed. The computer will then start displaying the results of your subject's task in the same form as shown in **Chapter III, Section 4, Data Display.** When the results have all been displayed, you will be returned to the function menu. If you wish, you can redisplay the results using F3 REDISPLAY CURRENT RESULTS.

F6 DELETE STORED RESULTS

The storage capacity of your computer is limited (**see Appendix B**). In order to make room for storing more subjects' results, you will delete those subjects whose data you no longer need. You start the process by pressing F6, to which the computer responds:

ENTER SUBJECT NAME>

The format of what you enter is the same as that previously described. The computer finds the subject and asks:

IS THE CORRECT SUBJECT (Y/N)?>

If you answer Y, you get one more chance to change your mind before the results are permanently erased. The computer asks:

DO YOU WANT TO DELETE THIS SUBJECT (Y/N)?

If you've changed your mind and answer N, you are returned to the function menu. If you answer Y, the computer will pause for a minute while it deletes the subject's test results. After it says:

SUBJECT DELETED

PRESS ENTER KEY TO CONTINUE

the function menu reappears.

The disk cartridge or tape we supplied included Michael Smith's data. You may want to try recalling it to see how these functions work.

APPENDIX A: System Requirements

The Gordon Diagnositic System (GDS)

The Data Analysis Program (DAP) is designed to simplify the collection of data generated by the GDS. Data is transmitted from the GDS to the computer via the standard communication link, an RS-232 connector. A female RS-232 connector is located on the back of the GDS. It "appears" to the computer like a serial printer, except that the data goes into the computer instead of out.

The RS-232 Cable

The GDS end of the cable needs to be a male 25 pin standard RS-232 connector. The requirement for the computer end depends on the computer you are using. Usually it is either a male or female 25 pin connector, but some computers use round connectors with fewer pins. As only 7 of the 25 pins are actually used, the cable itself also sometimes has fewer than 25 conductors.

Many computers will be able to run the DAP. The computer must have a serial communications port and should have another port to connect to a printer. The minimum memory requirement is 32K. The program and data storage can be on disk, recording tape, magnetic bubble cartridge, or Random Access Memory (RAM) depending on your computer. The DAP is written in Microsoft* or Applesoft* Basic, so you must have the appropriate one available on your machine. We cannot provide it with the DAP.

[*Microsoft and Applesoft are registered trademarks of Microsoft, Inc. and Apple Computers, Inc., respectively.]

The Printer

The printer is not essential to running the DAP, but a printer should be used to get the greatest benefit from the DAP. To provide graphic representations of the data on a variety of printers, the graphics are generated as a series of printed characters.

APPENDIX B: Specific Computers

Apple II

RS-232 port:	Super Serial Card & others
RS-232 cable:	25 pin, male/male
Storage:	floppy disk
Capacity (Subjects):	25
Startup:	automatic
Capital Letters:	automatic
Function Keys:	no
ENTER Key:	RETURN
BREAK Key:	CTRL & C
ESCAPE Key:	ESC

Sharp PC 5000

RS-232 port:	built in
RS-232 cable:	male/male
Storage:	magnetic bubble
Capacity (Subjects)	25
Startup:	automatic
Capital Letters:	Press CAPS key
Function Keys:	yes
ENTER Key:	arrow down, then left
BREAK Key:	ON/BRK
ESCAPE Key:	ESC

Radio Shack Model 100

RS-232 port:	built in
RS-232 cable:	25 pin, male/male
Storage:	tape cassette, computer memory
Capacity (Subjects):	6
Startup:	After loading from cassette, type RUN <ENTER>. Restarts automatically until program reloaded from tape.
Capital Letters:	press CAPS LOCK
Function Keys:	yes
ENTER Key:	ENTER
BREAK Key:	SHIFT & BREAK/PAUSE
ESCAPE Key:	ESC

IBM PC

RS-232 port:	Quadram and others
RS-232 cable:	25 pin, male/female
Storage:	floppy disk
Capacity (Subjects):	50
Startup:	automatic
Capital Letters:	press Caps Lock
Function Keys:	yes
ENTER Key:	arrow down, then left
BREAK Key:	Ctrl & Scroll Lock/Break
ESCAPE Key:	Esc

APPENDIX C: Organization of Data Files

This section describes the way the DAP stores its data. It is provided for the DAP user who wishes to write another computer program accessing the data.

Filename: TMONIT.DO
Type: Random Access (Serial in Model 100)
Record Length: 180 bytes
# of Records: 8 x maximum number of subjects stored
Record Format: Delay Task — 4 records

- Records 1-4 for Blocks 1-4
  - Up to 60 three digit numbers/block each number is the tenths of a second from the previous button press

Vigilance Task — 4 records

- Record 1 = 180 single digits for sequence of digits displayed
- Records 2-4 for Blocks 1-3
  - 180 spaces or hyphens each of whose positions corresponds to the character presented before a button press

In the Model 100, TMONIT.DO has been divided into DMONIT.DO and VMONIT.DO for the Delay and Vigilance tasks, respectively.

Filename: SUBJ.DO
Type: Random Access (Serial for Model 100)
Record Length: 220 bytes
# of Records: same as maximum number of subjects stored
Record Format:

CONTENT	bytes
Subject Identification	72
Subject Name	25
Space	1
Sex	1
Space	1
Birth	8
Examiner Name	13
Space	1
Comment	13
Space	1
Test Date	8
Delay Task Summary Statistics	72
1-10 Data Selector Displays each as 3 digits + space	40
Block Intervals each as 2 spaces + 2 digits	16
Block Lengths each as space + 3 digits	16
Delay Task Identifier = D	2
Vigilance Task Summary Statistics Same format as Delay Task	72
Vigilance Task Identifier = V9 or V1	2

APPENDIX D: Making Copies

All microcomputers and their storage media are subject to occasional failure. As a result, it is virtually essential to make a "backup" copy of anything important. The backup usually has the same form as your storage medium. For example, if you are using floppy disks, your backup will be a floppy disk which contains the exact same information as your original.

To provide for this at the outset, you should follow the directions supplied with your computer manual to copy the original DAP and put it away in a safe place. Use the copy to operate your computer. If the operating one fails, recopy the original DAP onto the cassette, disk, or cartridge you use every day.

If you are collecting a large amount of data on your storage medium, the same warning about making a backup copy is applicable. You may want to purchase additional cassettes, disks or cartridges to make backup copies of your stored data. All three storage media can be reused repeatedly when you no longer need the data you have saved on them.

ADDITIONAL INFORMATION:

The Gordon Diagnostic System may be used for the clinical assessment of stimulant drug response in the treatment of children with Attention Deficit Disorder (ADD). With the next publication of the instruction manuals, additional instructions (draft attached) will be provided for its use in this regard.

The foundation for the use of the GDS data used in this regard is based on research work performed by Dr. Russell Barkley and Dr. Mark Rapport and presented to the American Psychological Association Meeting in August 1985.

Russell Barkley presented data on a multimethod clinical protocol for the clinical assessment of stimulant drug responding. Barkley's approach involves a triple-blind, placebo controlled crossover design using two doses of methylphenidate (0.3 mg/kg and 0.5 mg/kg) and a placebo. To assess drug response, he uses parent and teacher ratings of ADD behaviors, problem settings and drug side effects; the GDS Tasks; and clinic playroom observations of ADD behavior during academic performance. Most of the measures in the protocol were sensitive to the effects of stimulant medication at both doses. The Commission score was found sensitive at the higher dose. Barkley urged that this objective, multimethod assessment protocol replace the current practice of relying upon subjective impression.

Mark Rapport from the University of Rhode Island presented data from an extensive drug response study that incorporated school-based observations of behavior. Analyzed on a group basis, most of the variables in the protocol (i.e. percent on task, percent work complete, percent work correct, teacher rating scales of attention, GDS Vigilance Task Commission Score) there were significant linear relationships with dosage levels of medication. There were also significant differences on these measures between each level of medication. In the analyses of group trends, the Vigilance Task Ommission score did not reach significance as a variable responsive to medication (the Delay Task was not used in this study). Subject-by-subject analyses, however, indicated that the failure of the score to reach significance was related to the idosyncratic response of children to the various dosage levels. In other words, a particular child might have improved his score significantly at one dose, but not at another - a phenomenon very much present, but to a lesser degree for the other variables. Rapport viewed these data as pointing up the need to judge children's drug response on an individual basis using a comprehensive battery of tests.

MONITORING OF PHARMACOTHERAPY

(DRAFT)

The Gordon Diagnostic System (GDS) can be used as part of an evaluation battery for the monitoring of pharmacotherapy. While the GDS is useful in this process, it is important that decisions regarding dosage levels include a comprehensive series of other assessment techniques.

Monitoring of pharmacotherapy is done optimally by testing the child when s/he is receiving several dosages of medication or a placebo administered in a random order. Only the pharmacist should know whether a child is on placebo or one of the dosage levels of medication. The following procedure is suggested:

- The child should be administered a dosage (placebo or otherwise). After 5-7 days the child should be tested with the GDS tasks.
- Following the GDS testing, the next dosage should be administered and, 5-7 days later, the GDS testing should be repeated.
- This procedure should be repeated for each drug condition.
- At the end of the trials, scores for each of the GDS tasks should be tabulated. A graph can be constructed which presents the GDS scores at each dosage level.

INTEPRETATION OF GDS SCORES FOR ANALYSIS OF A DRUG EFFECT

Research to date suggests that the Total Commission score is the most sensitive to the effects of stimulant medication. As such, improvement in this score from placebo condition to drug condition would indicate improvement in a child's ability to sustain attention. A general rule of thumb is to consider a score improved if the child's score changes from one range to another according to the Threshold Tables. For example, a child's score is considered significantly improved if, for example, the score moves from the Abnormal Range to the Borderline Range (or Borderline to Normal). The effects of various dosage levels can also be gauged with this method.

While current studies have been most supportive of the Commission score, it is suggested that all the scores be analyzed for changes related to medication status. It is again important, to emphasize that decisions surrounding the monitoring of pharmacotherapy should be based only on a comprehensive set of clinical data.

Journal of Abnormal Child Psychology, Vol. 7, No. 3, 1979, pp. 317-326

The Assessment of Impulsivity and Mediating Behaviors in Hyperactive and Nonhyperactive Boys[1]

Michael Gordon[2]
The Ohio State University

Twenty boys (6-8 years) rated by their teachers as hyperactive and a matched sample of nonhyperactive boys performed a task that required them to withhold responding for a set time interval in order to be rewarded (DRL 6-second schedule). Half of each group worked on a one-button console while the other half was provided with additional collateral buttons. Results indicated that hyperactive children were relatively unable to perform efficiently on the task, and that this deficit endured regardless of age, IQ, or experimental condition. DRL was thus found to discriminate accurately between teacher-rated and parent-rated hyperactive and nonhyperactive children. Furthermore, a wide variety of self-generated mediating behaviors was observed, and it was determined that a child's DRL performance was related to the kind of mediating behaviors he displayed. Results are discussed in terms of the clinical assessment of hyperactivity and the training of impulsive children.

Although the traditional approach has been to regard hyperactivity as a disorder of activity level, there is little evidence that the majority of children diagnosed as hyperactive are any more active, on the average, than a matched sample of nonhyperactive children, regardless of how activity is measured (Davis, Sprague, & Werry, 1969; Doubros & Daniels, 1966; Hutt, Jackson, & Level, 1966; Johnson, 1972; Lee & Hutt, 1964; McConnel, Cromwell, & Bialer, 1964; Ounsted, 1955; Schulman & Reisman, 1959). Moreover, viewing hyperkinesis solely in terms

Manuscript received in final form December 15, 1978.

[1] This report is based upon a dissertation submitted by the author in partial fulfillment for the Ph.D. degree. The efforts of Malcolm Helper, David Hothersall, Donald R. Meyer, Marilyn Moody, and Charles Wenar are gratefully acknowledged. The investigation was conducted while the author was a National Research Fellow, supported by Public Health Service Training Grant No. 5 T32 MH14608-02 to The Ohio State University.

[2] Address all correspondence to Michael Gordon, Upstate Medical Center, Division of Child and Adolescent Psychiatry, 750 East Adams Street, Syracuse, New York 13210.

0091-0627/79/0900-0317$03.00/0

of activity makes it difficult to understand many of the other symptoms that are often associated with this disorder, e.g., antisocial behavior, learning disabilities, distractibility, and excitability. The disenchantment with focusing on the quantity of a hyperactive child's activity has led some investigators to conceptualize hyperkinesis as a disturbance of impulse control and sustained attention. According to Douglas (1972), what typifies the hyperactive child is his inability to "stop, look and listen." She reviewed a body of research that appears to suggest that what characterizes the hyperactive child's behavior is not necessarily the sheer amount of activity, but its poor organization and control.

While the supporting studies cited by Douglas (1972) are intriguing, there is still need for a more straightforward assessment of the hyperactive's impulsivity. There is a tested technique, derived from research in operant conditioning, that provides an excellent opportunity to measure the hyperactive child's ability to inhibit behavioral responding. In this procedure, termed Differential Reinforcement of Low Rate Responding (DRL), responses that occur before a set time interval has passed are not reinforced and, moreover, they serve to reset the timer governing reinforcement. Thus a subject performing on a DRL 6-second schedule must wait at least 6 seconds between responses in order to receive a reinforcement. If a response is made before the 6 seconds have elapsed, the timer will reset and another 6 seconds must pass before a successful response can be made.

Normal laboratory animals performing on this task are able to develop efficient response patterns and typically come to be reinforced for 70% to 80% of the responses they make (Hothersall, Alexander, & Slonaker, 1972). As for human DRL performance, Weisberg and Tragakis (1967) demonstrated that even young children, rating from 15 to 40 months, were able to maintain low levels of responding. Latency-age boys (Stein & Landis, 1973) and college students (Randolph, 1964) have also been shown to perform efficiently, although the samples for most studies were extremely small. One major question posed by the present study is: How would a hyperactive child fare in a DRL situation in which behavioral responding must be suppressed? Might the DRL procedure provide an objective measure of impulsivity?

Researchers investigating DRL performances have also observed that subjects in a time-based situation will often engage in a sequence or chain of behaviors between reinforceable responses. The notion is that this collateral behavior sequence helps the subject "wait out" the required temporal delay between responses. Examples of such behaviors in animals include wood chewing and tail licking (Slonaker & Hothersall, 1972). Similarly, human subjects have become convinced that these extra mediating behaviors may set up the reinforced responses (Bruner & Revusky, 1961). Experimentally precluding such collateral behaviors has led to a disruption of the subject's ability to make temporal discriminations (Stein & Landis, 1973; Stein & Flanagan, 1974). The

frequency of extra movements made by hyperactive children during a delayed reaction time task was positively related to the efficiency of their performances (Douglas, 1972; Cohen, cited in Douglas, 1972), suggesting that collateral behaviors may play a role in helping these children gain impulse control. Another intent of the current study was to examine whether or not a situation that facilitates the development of collateral behaviors helps to improve the DRL performance of hyperactive and, for that matter, normal children. And, regardless of the availability of experimentally introduced opportunities for collateral responding, will these children spontaneously elicit mediating responses?

METHOD

Subjects

Twenty hyperactive boys between 6 and 8 years of age (72-97 months) and a matched sample of nonhyperactive boys participated in the study. The subjects were selected from the client roster of a child guidance clinic; they were all nonretarded and enrolled in school at an age-appropriate grade level. A child was considered hyperactive if his teacher rated him as such on the hyperactivity factor of the Behavior Rating Scale (Conners, 1969). In accordance with the findings of a normative study (Werry, Sprague, & Cohen, 1975), those children with a total score of at least 15 on this scale were allotted to the hyperactive group, while subjects receiving a score of less than 15 made up the control group. The two groups were statistically comparable for age.

Design

The experiment essentially followed the design and methodology presented by Stein and Landis (1975). Four equal-sized groups of 10 subjects each were involved: two groups of hyperactive and two groups of nonhyperactive boys. One hyperactive and one nonhyperactive group performed on a DRL 6-second schedule without the availability of collateral buttons (DRL ONLY), while the other two groups worked on DRL using the five-button console (DRL + COLLATERAL).

Apparatus

Each subject was seated before a console containing either one (DRL ONLY) or five (DRL + COLLATERAL) microswitch buttons mounted 4.0 cm

apart. On a separate console, positioned immediately behind the button box, was a chassis containing a large six-digit add counter, which indicated the cumulative number of reinforcements (points) earned, and a red light, which lit for 2 seconds upon completion of a successful response. For the DRL ONLY condition, the one red button was programmed to produce reinforcement according to a DRL 6-second schedule. In the DRL + COLLATERAL condition, only one red button was programmed with the DRL schedule. The remaining four black collateral keys did not yield reinforcements, but they were individually monitored for the number of times they were pressed, and the response pattern was retained by an Esterline Angus Event Recorder. All sessions were videotaped in their entirety. The programming equipment, counters, timers, recorders, and monitoring equipment were housed in a separate room, which was joined to the experimental room by a one-way mirror and an intercom system.

Procedure

Prior to the testing session, the mother of every subject received an appropriate introduction to the study. The child was then taken individually to the experimental room, which contained only a table, chairs, and the testing apparatus. He then received the following instructions (adapted from Stein & Landis, 1975):

> You're going to play a game in which you will get a chance to win a lot of M&M's. Do you see this light [pointing to the reward indicator on the reinforcement box]? Every time you make this light go on you'll earn an M&M, and this counter [pointing] will keep track of how many M&M's you've won. At the end of the game I'll give you all the M&M's that you earned. Now, to make the light go on all you have to do is push this red button [pointing], and wait a little while before pressing it again. You just press this red button, wait a while, then press it again. If you press it again too soon, though, then you will not get a point, the light won't go on, and you'll have to wait a while before you can press it to get another point. But if you push the button, wait a while, then push it again, you'll get a point every time.

For the DRL + COLLATERAL groups, these instructions were added:

> These four black buttons will not give you points if you press them but, if you want, you can play with them while you're waiting to press the red button again.

The general instructions continued:

> Do you have any questions? [All questions were answered.] OK, I'll leave the room now and turn on the machine. When you hear my voice coming over the speaker saying, "You can start now," then you can begin playing the game. I'll come back when the game is over and bring you your M&M's. Have fun!

Each subject had a total of 15 minutes to earn points on the DRL schedule. At the session's end he was asked, "How did you do it?" and "Did you do

anything to help you wait?" Following his response, he was paid his due of M&M's, thanked for his participation, and returned to his mother.

Measures

The DRL task yields three measures of a subject's performance: the total number of responses, the number of reinforcements earned, and the efficiency score. Since this last score represents the percentage of reinforced responses, i.e., the number of reinforcements divided by the number of responses, it is considered to be the most sensitive indicator of a child's performance. Every subject's response, reinforcement, and efficiency scores were obtained for each of three 5-minute time blocks, and then totaled for the entire 15-minute session. Also monitored were the number of responses emitted on the collateral keys (if available) and the pattern in which they were pressed. Data on IQ, exact age, school grade, and parent ratings were also collected.

RESULTS

DRL Analysis

Analyses of variance (ANOVAs) were performed on the three operant measures, and the mean scores for each trial and for the whole session are shown in Figure 1. There was a highly significant main group (hyperactive vs. nonhyperactive) effect for all three DRL variables, especially the efficiency score, $F(1, 36) = 18.33$, $p < .0013$. Children rated as hyperactive by their teachers thus tended to emit more responses and receive fewer reinforcements during each of the time blocks than did children judged nonhyperactive. This main effect remains strong if the data are reanalyzed according to the parents' classifications of hyperactivity. The relative inefficiency of these hyperactive children suggests that they are significantly more impulsive as measured by the DRL 6-second schedule.

There were no Condition (DRL ONLY vs. DRL + COLLATERAL) effects on any of the dependent variables (efficiency score, $F(1, 36) = .133, p < .717$), so that whether a child played with the one-button or the five-button console had an insignificant effect upon his DRL performance. A few Group X Condition interactions did emerge: Hyperactive children in the DRL + COLLATERAL condition tended to make more responses than did their counterparts in the DRL ONLY condition, $F(1, 36) = 6.4$, $p < .016$, and nonhyperactive COLLATERAL subjects did significantly better according to the efficiency score than did the nonhyperactives in the DRL ONLY situation.

The ANOVAs yielded only one marginally significant main effect for Trials (time blocks): All groups tended to receive more rewards for their second-

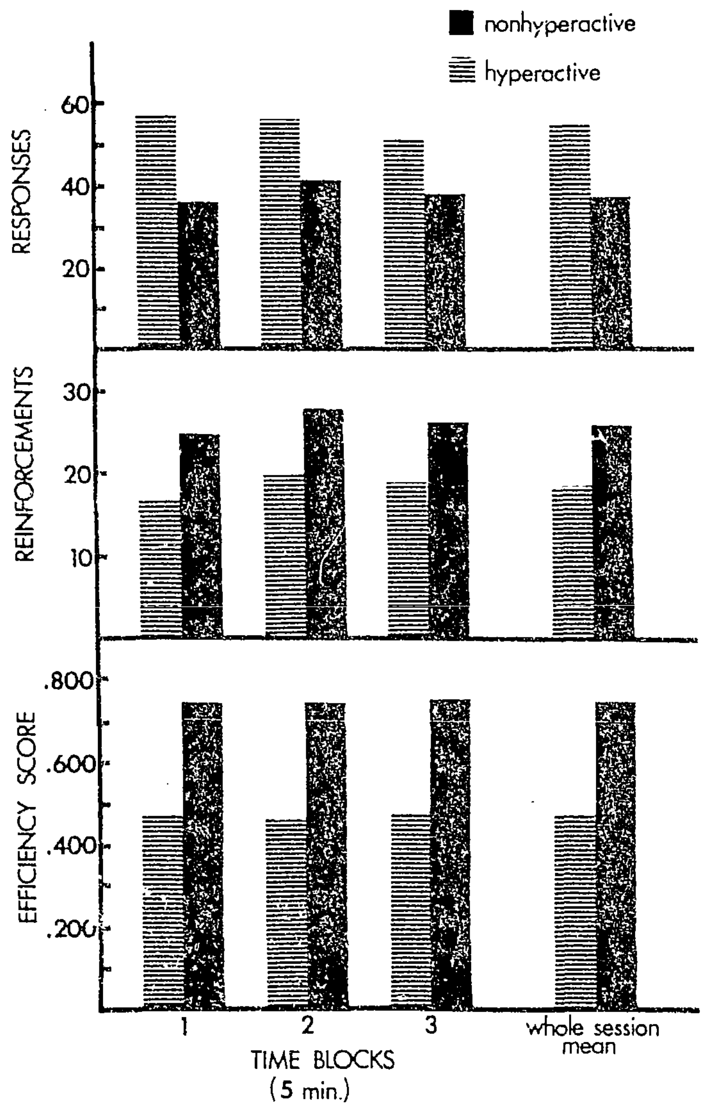


Fig. 1. The performance of hyperactive and nonhyperactive groups over three time blocks of a DRL 6-second schedule. Top: responses; middle: reinforcements; bottom: efficiency score.

trial performance, $F(2, 72) = 3.08$, $p < .051$. In general, the second trial was the most characteristic of a child's DRL work. For nearly all the various analyses, what held true for each trial taken separately, or for the session as a whole, usually held true for the second-trial measures and, again, particularly for the second time block efficiency score. This observation is supported by the intertrial correlations for each measure. Since the correlations between Trial 2 and Trial 3 tend to be quite high, mean $r = .69$, $p < .001$, it would suggest that a boy's performance was apt to stabilize by the second trial. When these interblock correlations are computed for each group separately, there is evidence that the DRL conduct of nonhyperactive subjects generally consolidated from one trial to the next as they locked into a relatively efficient pattern, mean $r = .75$, $p < .001$. The hyperactive children were more erratic across trials than the nonhyperactive children, mean $r = .54$, $p < .001$, although their performance was also relatively stable, as the intertrial correlations indicate.

In order to uncover relationships between the operant scores and a variety of descriptive data, all variables generated by this study were intercorrelated. It was first found that the subject's age is thoroughly unrelated to any other aspect of his performance or to the hyperactivity ratings supplied by the teacher or parent. The teacher ratings of hyperactivity using the Behavior Rating Scale (Conners, 1969) were significantly correlated with all DRL variables; their correlation with the whole-session efficiency score, for example, was $r = .53$, $p < .01$. The parent rating was also tied to the major DRL scores (mean $r = .35$, $p < .05$), though slightly less than were the teacher ratings. The correlation between parents and teachers, themselves, was $r = .61$ – highly significant ($p < .001$), but not so high as to instill total confidence in their reliability. Parents and teachers did closely agree (34/40 or 85% of the cases) on the gross classification of a child as hyperactive or nonhyperactive. Of the six contesting cases, parents rated four teacher-classified nonhyperactives as hyperactive and two teacher-classified hyperactives as nonhyperactive. It seems, therefore, that parents were apt to classify more children as hyperactive. Yet the teachers' mean ratings of the hyperactive boys (22.4) was slightly higher than those rendered by parents (20.0). So, teachers classified fewer children as hyperactive, but they judged the hyperactives as more severely hyperactive.

Analysis of Mediating Behaviors

As has been the case for almost all animal and human subjects working on a DRL schedule, the children who participated in this study tended to engage in a variety of interresponse behaviors that were either physically manifested or were reported to the examiner in response to a standard, end-of-session question: "How did you do it?" or "Did you do anything to help you wait?" A list of some of the mediating behaviors that were observed appears in Table

Table I. A List of Observed Mediating Behaviors

Circling DRL button with finger 9 times
Swinging legs 10, 12, or 20 times
Counting with lips
Counting out loud numbers or ABC's
Blowing on reward box
Singing out loud
Shaking reward box 10 times
Hitting knee with right hand 20 times
Foot-tapping 16 times
Tapping finger 10 times on button box
"Walking" fingers around DRL button 9 times
Stomping with foot 9 or 10 times
Running around table 1 time
Hitting side of box
Jumping jacks 4 times
Hitting collateral buttons

I. (It should be noted that only 5 of the 20 Collateral group used the collateral buttons for their interresponse regimens.) In order to be considered a mediating behavior, the sequence had to be repeated among at least 10 responses within a particular trial. Judgments concerning collateral responding were made blind from the videotapes. As it turned out, every subject engaged in, or reported, some mediating behavior. Ninety percent of the hyperactive subjects engaged in an observable mediating behavior, compared to 45% of the nonhyperactives. On the other hand, just 30% of the hyperactive boys reported the use of a nonbehavioral mediator, as opposed to 80% of the nonhyperactive subjects. A roughly comparable percentage of each group used both approaches (hyperactives = 20%, nonhyperactives = 25%).

There was an inverse correlation between the use of a physical collateral and both the efficiency ($r = -.32$, $p < .05$) and reinforcement ($r = -.55$, $p < .005$) scores, especially after the first trial. Hence, children who used an observable collateral scheme did worse on the DRL schedule; they were also apt to be rated as more hyperactive by their parents and teachers. The opposite pattern emerged for cognitive mediators as there was a very high correlation between the children's reports of counting and all measures for each trial, particularly blocks 2 and 3, mean $r = .64$, $p < .001$. Thus subjects who engaged in cognitive collaterals fared better on DRL and were usually judged less hyperactive than their noncounting counterparts. Boys who used both strategies during the session also had better scores, though not as high as the scores of those who only counted.

DISCUSSION

The prediction that children classified as hyperactive would encounter greater difficulties in managing the DRL task than would nonhyperactives was confirmed by the results of this study. According to every indicator of DRL performance, the hyperactive boys were relatively unable to refrain from making a high percentage of incorrect responses. Consequently, these results lend further credence to Douglas's (1972) proposal of an impulsivity dimension underlying the hyperkinetic syndrome, and to the validity of other measures of impulsivity presented by Sykes (cited in Douglas, 1972) and Campbell (cited in Douglas, 1972).

The experimental Condition (DRL ONLY vs. DRL + COLLATERAL) failed to emerge as a significant source of variance. However, given the observation that children engaged in many self-generated interresponse behaviors, this finding is not at all surprising. The fact is that a child will develop some type of interresponse strategy whether or not the experimenter provides a few extra buttons or telegraph keys. Indeed, in their attempts to quantify collateral responding, it appears that past investigators have missed the rich bounty of self-elicited behaviors that children exhibit during the course of a DRL session. Aside from their diversity, these schedule-induced behaviors are fascinating for the rigid manner in which they are executed; once a child chanced upon a successful strategy, he usually stuck with it ardently.

That the proclivity to engage in a physical or cognitive mediating behavior was so closely related to DRL performance as well as to ratings of hyperactivity is a highly intriguing result. Even with a relatively gross coding system, a strong pattern surfaced in which boys who displayed physical collaterals did worse on DRL and were rated more hyperactive than were boys who used cognitive means. In addition, many more hyperactive than nonhyperactive boys generated physical collaterals, an observation that may support the notion that a certain segment of the hyperactive child's hyperactivity is actually a series of mediating behaviors developed in an effort to control his impulses.

While it appears certain that cognitive mediators are more efficient strategies than physical ones, it is impossible to determine from these data whether physical collaterals actually hindered DRL efficiency or simply failed to facilitate it as much as the cognitive types did. The only hint available is the finding that the 18 subjects who employed physical collaterals during several trials had significantly higher efficiency scores than did the 9 children whose physical mediating was limited to one time block, $p < .05$. The precise nature of the interaction between a physical collateral and the DRL performance of hyperactive children is a very important avenue of research that warrants further exploration.

An unexpected aspect of these results was the extent to which the DRL variables could accurately discriminate between groups. According to the inter-trial correlations, the efficiency score can serve as a reliable and stable predictor of teacher-classified hyperactivity. Consequently, it is possible that this task could be developed into a useful clinical tool that could aid the diagnostician in the assessment of impulsivity and the classification of hyperactivity.

REFERENCES

Bruner, A., & Revusky, S. A. Collateral behavior in humans. *Journal of the Experimental Analysis of Behavior*, 1961, *4*, 349-350.

Cantwell, D. P. *The hyperactive child: Diagnosis, management, current research.* New York: Spectrum Publications, 1975.

Conners, C. K. A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry*, 1969, *126*, 152-156.

Davis, K., Sprague, R., & Werry, J. Stereotyped behavior and activity level in severe retardates: The effects of drugs. *American Journal of Mental Deficiency*, 1969, *72*, 721-727.

Doubros, S., & Daniels, G. An experimental approach to the reduction of overactive behavior. *Behaviour Research and Therapy*, 1966, *4*, 251-258.

Douglas, V. I. Stop, look and listen: The problem of sustained attention and impulse control in hyperactive and normal children. *Canadian Journal of Behavioral Science*, 1972, *4*, 259-282.

Gittis, A., & Hothersall, D. DRL performance of juvenile rats with septal lesions. *Physiological Psychology*, 1974, *2*, 38-42.

Hothersall, D., Alexander, D., & Slonaker, R. The DRL deficit of rats with septal lesions: Effects of extended training in a mediated environment. *Psychonomic Science*, 1972, *29*(1), 34-36.

Hutt, D., Jackson, P., & Level, M. Behavior parameters and drug effects: A study of a hyperkinetic epileptic child. *Epilepsia*, 1966, *7*, 250-259.

Johnson, C. F. Limits on the measurement of activity levels in children using ultrasound and photoelectric cells. *American Journal of Mental Deficiency*, 1972, *77*, 301-310.

Lee, D., & Hutt, C. A play-room designed for filming children: A note. *Journal of Child Psychology and Psychiatry*, 1964, *5*, 263-265.

McConnel, T. B., Cromwell, R. F., & Bialev, I. A. Studies in activity level: VIII. *American Journal of Mental Deficiency*, 1964, *68*, 647-651.

Ounsted, C. The hyperactive syndrome in epileptic children. *Lancet*, 1955, *269*, 303-311.

Randolph, J. J. A further examination of collateral behavior in humans. *Psychonomic Science*, 1964, *3*, 227-228.

Schulman, J. L., & Reisman, J. M. An objective measure of hyperactivity. *American Journal of Mental Deficiency*, 1959, *64*, 455-456.

Slonaker, R. L. & Hothersall, D. Collateral behaviors and the DRL deficit of rats with septal lesions. *Journal of Comparative and Physiological Psychology*, 1972, *80*, 91-96.

Stein, N., & Flanagan, S. Human DRL performance, collateral behavior and verbalization of the reinforcement contingency. *Bulletin of the Psychonomic Society*, 1974, *3*, 27-28.

Stein, N., & Landis, R. Mediating role of human collateral behavior during a spaced responding schedule of reinforcement. *Journal of Experimental Psychology*, 1973, *97*, 28-33.

Stein, N., & Landis, R. Differential reinforcement of low rates performance by impulsive and reflective children. *Journal of Experimental Child Psychology*, 1975, *19*, 37-50.

Weisberg, P., & Tragakis, C. J. Analyses of DRL behavior in young children. *Psychological Report*, 1967, *21*, 709-715.

Werry, J. S., Sprague, R. L., & Cohen, M. N. Conners' Teacher Rating Scale for use in drug studies with children: An empirical study. *Journal of Abnormal Child Psychology*, 1975, 3(3), 217-229.

The Objective Assessment of Attention Deficit Disorders

Michael Gordon, Ph.D.
State University of New York, Upstate Medical Center

F. Daniel McClure, Ph.D.
David C. Wilson Neuropsychiatric Hospital

Objective and reliable techniques for the assessment of Attention Deficit Disorders (ADD) with and without hyperactivity in children have remained largely unavailable to psychologists, educators and pediatricians. As a consequence, they have tended to base their evaluations of ADD solely on observations and teacher reports, or on measures such as the Wechsler scales, which have been shown to discriminate poorly between groups of hyperactive and nonhyperactive children. Over the past 6 years we have been developing a behavioral measure of ADD, called the Gordon Diagnostic System (GDS), which has shown considerable promise as a precise, valid and efficient technique for the diagnosis of attention disorders. The GDS contains two measures: the Delay Task, which measures the ability to inhibit responding; and the Vigilance Task, which assesses sustained attention. In our research studies, we have shown that these game-like tasks differentiate accurately between hyperactive and nonhyperactive children. Normative data on GDS performance have been established based on a study of 220 nonhyperactive boys and girls from 6 through 11 years of age. Information on test-retest reliability is also available.

INTRODUCTION

Clinicians, educators, and physicians have been in chronic need of reliable and valid techniques for the assessment of Attention Deficit Disorders with and without hyperactivity (ADD) and impulsivity in children. The instruments that are currently available for the diagnosis of ADD fall into three basic categories: those actually intended for the measurement of other aspects of behavior which have been adapted to assess impulsivity, those which are largely limited to laboratory investigations, and questionnaires.

Paper presented at the 91st Annual Convention of the American Psychological Association at Anaheim, CA, August, 1983

Correspondence address: 750 East Adams Street
Syracuse, New York 13210

Most prominent among the instruments originally designed for the clinical assessment of other behavioral domains is the Wechsler Intelligence Scale for Children-Revised (WISC-R). While selected subtests or groups of subtests from the WISC-R are widely used in the identification of hyperactivity, studies have shown that they do not discriminate accurately between groups of hyperactive and nonhyperactive children (Douglas, 1972). Moreover, those subtests assess a wide range of abilities and traits that are largely unrelated to impulsivity (Douglas, 1972). Thus, while all children with ADD may perform poorly on these measures, there are many nonhyperactive children (e.g. learning disabled, or those with visual-motor deficits) who may also have difficulties with them for reasons other than impulsivity.

Among those instruments best suited for laboratory investigations, the most effective in differentiating hyperactive from nonhyperactive children are the Continuous Performance Test (CPT; Rosvold & Mirsky, 1956), Delayed Reaction Time Test (DRTT; Cohen, Douglas & Morganstern, 1972), and the Matching Familiar Figures Test (Kagan, 1966). While the CPT and DRTT have been demonstrated to be excellent diagnostic measures, their technological sophistication and prohibitive expense make them unsuitable for widespread clinical use. The Matching Familiar Figures Test is quick, easy to administer, and has been demonstrated to be diagnostically sound. However, its reliability and validity have been the subject of considerable controversy which persists despite revisions of the measure (Egeland & Weinberg, 1976).

There are a variety of other measures available such as the Gardner Steadiness Tester (Gardner, 1979), Human Figure Drawings (Machover, 1949), certain configurations of Rorschach indices, and paper-and-pencil "vigilance" tasks. As a group, these tend to suffer drawbacks stemming from impracticality of administration, limited yield of information, or inadequate standardization.

Behavioral inventories or questionnaires have also enjoyed widespread use, and many, such as the Teacher Rating Scale (Conners, 1969), the Child Behavior Problem Check List (Achenbach, 1978), and the Self-Control Rating Scale (Kendall & Wilcox, 1979), have demonstrated good discriminative validity. However, there are serious inherent drawbacks to such a format, including rater bias, the possibility of "halo" effects, item overlap, vague rating criteria, issues of basic validity, and the fact that questionnaires yield no direct behavioral data.

There is thus an obvious and compelling need for a more straightforward, clinically applicable, and precise measure of impulsivity in the child with ADD. The Gordon Diagnostic System (GDS) shows considerable promise as a device which can help fill this gap in clinical service delivery. The GDS is an easily-administered, game-like task which measures levels of impulsivity without interference from other factors such as intelligence or visual-motor skills. It provides objective data on a particular child's ability to inhibit behavioral responding. Through observation of the child's performance and from responses to several post-test inquiries, important insights can be gained into a child's problem-solving style. This information can have strong bearing on treatment recommendations and educational planning. Thus, the GDS is unique because it yields not only quantitative data, but also a wealth of pertinent behavioral information which is crucial to the diagnostic process.

DESCRIPTION OF THE GDS

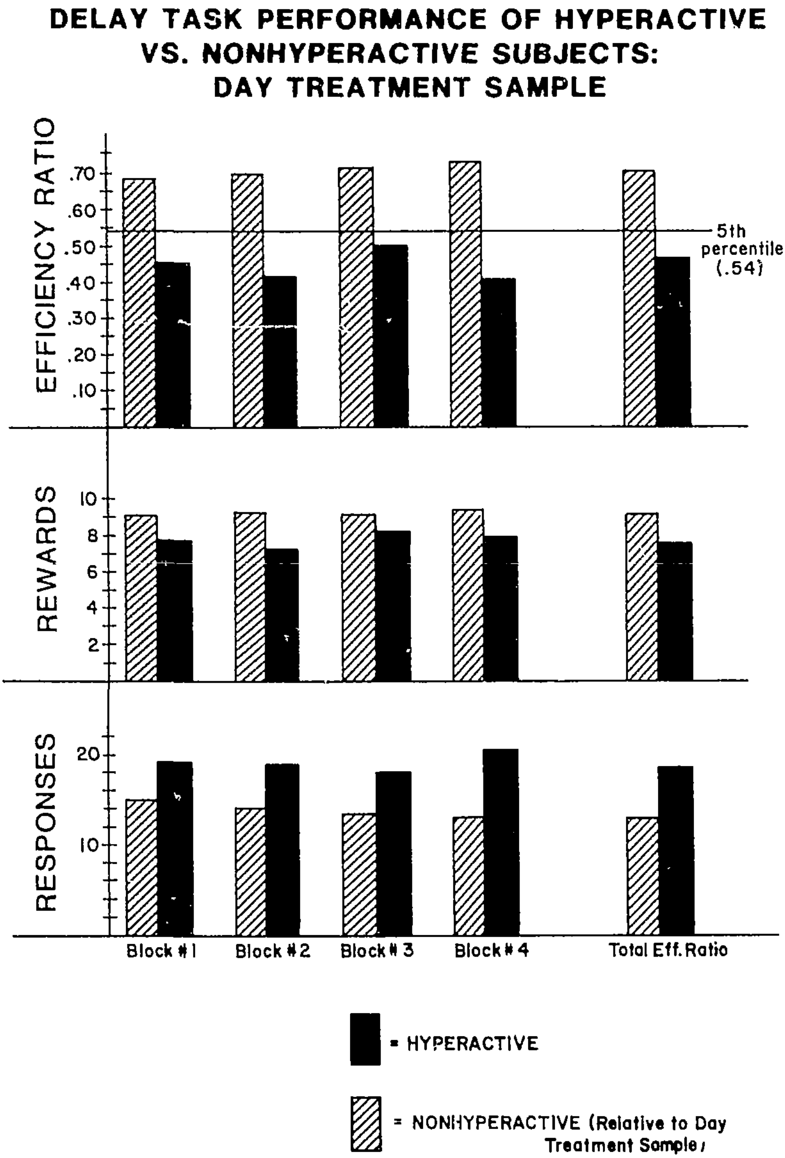
The GDS is a portable, electronic unit which allows for the administration of two tasks. The Delay Task is based upon a Differential Reinforcement of Low Rates (DRL) behavior schedule and requires the child to inhibit responding in order to gain a reward. Specifically, the child is instructed to press a button, wait, and then press the button again. If he/she refrains from responding for at least six seconds, a light flashes and a reward counter increments. If the child responds before the interval elapses then the timer resets and no reward points are recorded. The GDS Delay Task generates three major scores: the number of responses (button presses), the number of correct responses (Rewards) and the Efficiency Ratio which represents the percentage of correct responses.

The GDS also contains the Vigilance Task which is a version of the Continuous Performance Test (Rosvold, et al., 1956). The child is presented with a series of digits on the front display and is told to press the button every time a "9" appears that had been preceded by a "1". The GDS records the number of correct presses, the number of times the child fails to press the button when the "1/9" combination appears (errors of Omission), and the number of extraneous button presses (errors of Commission). For testing younger children the GDS contains a "1" mode, which requires the subject to press the button every time a "1" appears. The same performance measures are recorded.

RESEARCH STUDIES WITH THE GDS

Data concerning the GDS Delay Task come from six studies conducted over as many years. The original experiment (Gordon, 1979), performed on an early version of the GDS, compared 20 hyperactive and 20 nonhyperactive children selected from the patient roster of a child guidance clinic. There were highly significant differences between the two groups for all three Delay Task variables: Responses, Rewards and the Efficiency Ratio (Rewards/Responses). The children classified as hyperactive had profound difficulties in dealing successfully with the delay inherent in the task.

The original finding was recently replicated (McClure & Gordon, in press) with patients placed in a day treatment program for seriously disturbed youngsters. Differences in Delay Task efficiency ratios between the 20 children classified as hyperactive and the 20 nonhyperactive ones in the comparison group were strikingly significant ($p < .0001$).


DELAY TASK PERFORMANCE OF HYPERACTIVE VS. NONHYPERACTIVE SUBJECTS: DAY TREATMENT SAMPLE
EFFICIENCY RATIO
.70
.60
.50
.40
.30
.20
.10
5th percentile (.54)
REWARDS
10
8
6
4
2
RESPONSES
20
10
Block #1
Block #2
Block #3
Block #4
Total Eff. Ratio
= HYPERACTIVE
= NONHYPERACTIVE (Relative to Day Treatment Sample)

To further judge the discriminative ability of the GDS *post hoc* analyses were performed with groups constituted on the basis of being identified as hyperactive or nonhyperactive by four criterion measures. These included the Teacher Rating Scale (Conners, 1969), the Self-Control Rating Scale (Kendall & Wilcox, 1979), the Teacher Report Form (Achenbach, 1980), and the Matching Familiar Figures Test - 20 (Cairns & Cammock, 1978). For subjects who were identified as hyperactive (n=8) or nonhyperactive (n=8) on the basis of all four criteria, there was agreement between classifications derived from the GDS Delay Task efficiency ratio and the criterion measures in 15 cases (94%, Fisher's Exact Test, p=.0007). When groups were classified as hyperactive or nonhyperactive by at least three of the four criterion measures, there was agreement in 29 of 32 cases (91%, Fisher's Exact Test, p=.000003).

DIAGNOSTIC ACCURACY OF THE DELAY TASK: DAY TREATMENT SAMPLE

Criterion Measures: Teacher Rating Scale (Connors, 1969)
Self-Control Rating Scale (Kendall & Wilcox, 1979)
Teacher Report Form (Achenbach, 1980)
Matching Familiar Figures Test (Cairns & Cammock, 1978)

Hyperactive Group: Identified as Hyperactive on at least 3 measures

Nonhyperactive Group: Identified as Nonhyperactive on at least 3 measures

	Efficiency Ratio High (>.55)	Efficiency Ratio Low (≤.55)
Hyperactive (n=15)	1	14
Nonhyperactive (n=17)	15	2

p=.000003 (Fisher Exact Test)

There was agreement between the Delay Task Efficiency Ratio and the criteria on 91% of the cases.

Two additional studies appear to establish the validity of the GDS as a measure of impulsive/hyperactive behavior. In a group of 31 nonhyperactive subjects, GDS performance was significantly correlated with factors derived from teacher ratings. It was also related to impulsivity as measured by the

Bender-Gestalt test and the Rorschach Inkblots. Analyses of data from the files of 32 nonhyperactive outpatient children also indicate strong relationships between GDS performance and parent ratings of impulsive behavior.

CORRELATIONS BETWEEN DELAY TASK TOTAL EFFICIENCY RATIO AND COLLATERAL MEASURES

Day Treatment Sample-Combined Hyperactive and Nonhyperactive	N	r
Age	40	.15
IQ	40	.01
Matching Familiar Figures - Latency	40	.46**
Matching Familiar Figures - Errors	40	-.36*
Self Control Rating Scale (Kendall, 1979)	40	-.16
Child Behavior Checklist Hyperactivity Factor (Achenbach, 1970)	40	.33*
Behavior Rating Scale	40	-.58***

Outpatient Clinic Sample-Nonhyperactive Subjects	N	r
Behavior Rating Scale (Conners, 1969) Hyperactivity Factor:		
Completed by Teachers	30	-.53**
Completed by Parents	30	-.50**
Child Behavior Checklist Hyperactivity Factor (Achenbach, 1970)	27	-.44**
Bender-Gestalt Impulsivity Score	29	.42**
Rorschach Human Movement (M)	31	.68**

*p <.05
**p <.01
***p <.001

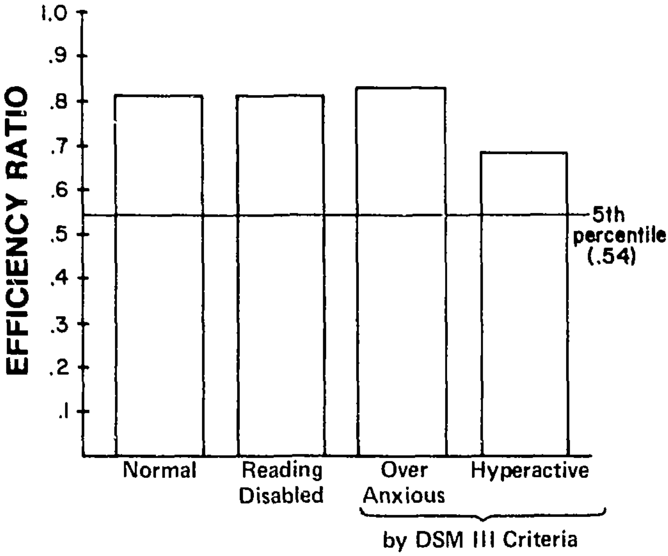
To establish norms for the GDS Delay Task scores, a standardization study was recently completed in the public schools. The sample included a total of 192 boys and girls in three age groups (6-7 yrs.; 8-9 yrs.; 10-11 yrs.). It was determined that the mean Efficiency Ratio for the Delay Task (Total Rewards/Total Responses) was .84 (SD = .13) and that the scores were unaffected by age, sex, IQ and socioeconomic status. As with previous studies, performance remained consistent throughout the 8-minute session. To establish test-retest reliability, half the group were retested after 30 - 45 days, and it was found that the scores remained stable over that period. Norms are also available for the Vigilance Task of the GDS for a sample of 92 boys and girls.

**MEANS OF GDS VARIABLES:
STANDARDIZATION SAMPLE**

	6-7 year olds				8-9 year olds				10-11 year olds			
	Boys		Girls		Boys		Girls		Boys		Girls	
lay Task (n = 192)	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Total Rewards	40.0	9.0	36.6	12.0	42.8	10.5	42.2	10.4	47.3	10.5	45.9	8.4
Total Responses	52.6	15.2	47.1	17.7	53.2	14.9	52.0	14.2	54.4	13.7	53.8	9.8
Total Efficiency Ratio	.80	.16	.83	.18	.83	.13	.83	.13	.88	.09	.86	.11
Vigilance Task (n = 91)	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Total Correct	30.7	9.4	25.7	11.8	37.6	6.0	38.5	3.7	41.8	1.7	41.5	2.3
Total Omissions	14.3	9.4	19.3	11.8	7.4	6.0	6.5	3.7	3.2	1.7	3.5	2.3
Total Commissions	16.1	15.5	9.3	5.0	11.8	10.7	9.6	10.3	6.9	8.0	5.1	4.1

Also completed recently was a study of 60 children who were classified by the school district as falling into one of four groups: Attention Deficit Disorder, reading disabled, overanxious and normal. The GDS Delay and Vigilance Tasks successfully differentiated between the ADD group and the other three categories. It was also found that the correlations among the Delay Task Efficiency Ratio and the Vigilance Task scores fell in the .6 to .7 range of Pearson's correlation coefficient.

**MEAN EFFICIENCY RATIO:
SCHOOL REFERRED SAMPLE**



The hyperactive group had a lower mean Efficiency Ratio than the other 3 groups.

While the GDS has successfully differentiated between groups classified as hyperactive and nonhyperactive, the accuracy of the results is often difficult to assess because of the absence of reliable, independent criteria for making the classification of hyperactivity. Even when multiple criteria are employed, the degree of diagnostic agreement among the criterion measures tends to be poor. Consequently, when the GDS classifies a child differently, the other criteria may well be at fault. Indeed, when we have pursued such cases with follow-up interviews of parents and teachers, the GDS classification has, with few exceptions, seemed most plausible. (In one case, a girl classified as impulsive by the GDS had been classified nonhyperactive by the teacher's rating. When asked about the child, the teacher responded, "Oh, she's real hyper but she's so darling that I don't really think she's hyperactive.") Our experience to date has led us to orient future research away from criterion-based studies and towards those which focus on behavioral correlates of GDS performance.

The GDS will be further evaluated at 20 research sites in this country and Canada. In addition to expanding our data base, several projects will be investigating whether the device is sensitive to the effects of stimulant medication, while others will be examining the possibility of using it as a training procedure. There is also ongoing research into the possibility of including other tasks into the GDS which will enhance the yield of information.

REFERENCES

Achenbach, T. The child behavior profile. I. Boys aged 6-11. *Journal of Consulting and Clinical Psychology,* 1978, 46, 478-488.

Achenbach, T. and Edelbrock, C. The classification of children's psychopathology: A review and analysis of empirical efforts. *Psychological Bulletin,* 1978, 85, 1275-1301.

Cairns, E. and Cammock, T. Development of a more reliable version of the Matching Familiar Figures Test. *Developmental Psychology,* 1978, 14, 555-560.

Cohen, N., Douglas, V., & Morganstern, G. The effect of methylphenidate on attentive behavior and autonomic activity in hyperactive children. *Psychopharmacologia.* 1972, 22, 282-294.

Conners, C. A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry,* 1969, 126, 884-888.

Douglas, V. Stop, look, and listen: The problem of sustained attention and impulse control in hyperactive children. *Canadian Journal of Behavioral Sciences,* 1972, 4, 259-282.

Egeland, B. and Weinberg, R. The Matching Familiar Figures Test: A look at its psychometric credibility. *Child Development,* 1976, 47, 483-491.

Gardner, R. *The Objective Diagnosis of Minimal Brain Dysfunction.* Cresskill, N.J.: Creative Therapeutics, 1979.

Gordon, M. The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive boys. *Journal of Abnormal Psychology,* 1979, 7, 317-326.

Kagan, J. Reflectivity-impulsivity: The generality and dynamics of cognitive tempo. *Journal of Abnormal Psychology,* 1966, 71, 17-24.

Kendall, P. and Wilcox, L. Self-control in children: Development of a rating scale. *Journal of Consulting and Clinical Psychology,* 1979, 7, 317-326.

McClure, D.F. and Gordon, M. The performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. *Journal of Abnormal Child Psychology,* in press.

Machover, K. *Personality Projection in the Drawing of the Human Figure.* Springfield, Ill., Charles G. Thomas, 1949.

Rosvold, H., Mirsky, A., Sarason, I., Bransone, E., & Beck. A continuous performance test of brain damage. *Journal of Consulting Psychology.* 1956, 20, 343-352.

Journal of Abnormal Child Psychology, Vol. 7, No. 3, 1979, pp. 317–326

The Assessment of Impulsivity and Mediating Behaviors in Hyperactive and Nonhyperactive Boys[1]

Michael Gordon[2]
The Ohio State University

Twenty boys (6-8 years) rated by their teachers as hyperactive and a matched sample of nonhyperactive boys performed a task that required them to withhold responding for a set time interval in order to be rewarded (DRL 6-second schedule). Half of each group worked on a one-button console while the other half was provided with additional collateral buttons. Results indicated that hyperactive children were relatively unable to perform efficiently on the task, and that this deficit endured regardless of age, IQ, or experimental condition. DRL was thus found to discriminate accurately between teacher-rated and parent-rated hyperactive and nonhyperactive children. Furthermore, a wide variety of self-generated mediating behaviors was observed, and it was determined that a child's DRL performance was related to the kind of mediating behaviors he displayed. Results are discussed in terms of the clinical assessment of hyperactivity and the training of impulsive children.

Although the traditional approach has been to regard hyperactivity as a disorder of activity level, there is little evidence that the majority of children diagnosed as hyperactive are any more active, on the average, than a matched sample of nonhyperactive children, regardless of how activity is measured (Davis, Sprague, & Werry, 1969; Doubros & Daniels, 1966; Hutt, Jackson, & Level, 1966; Johnson, 1972; Lee & Hutt, 1964; McConnel, Cromwell, & Bialer, 1964; Ounsted, 1955; Schulman & Reisman, 1959). Moreover, viewing hyperkinesis solely in terms

Manuscript received in final form December 15, 1978.
[1] This report is based upon a dissertation submitted by the author in partial fulfillment for the Ph.D. degree. The efforts of Malcolm Helper, David Hothersall, Donald R. Meyer, Marilyn Moody, and Charles Wenar are gratefully acknowledged. The investigation was conducted while the author was a National Research Fellow, supported by Public Health Service Training Grant No. 5 T32 MH14608-02 to The Ohio State University.
[2] Address all correspondence to Michael Gordon, Upstate Medical Center, Division of Child and Adolescent Psychiatry, 750 East Adams Street, Syracuse, New York 13210.

0091-0627/79/0900-0317$03.00/0

of activity makes it difficult to understand many of the other symptoms that are often associated with this disorder, e.g., antisocial behavior, learning disabilities, distractibility, and excitability. The disenchantment with focusing on the quantity of a hyperactive child's activity has led some investigators to conceptualize hyperkinesis as a disturbance of impulse control and sustained attention. According to Douglas (1972), what typifies the hyperactive child is his inability to "stop, look and listen." She reviewed a body of research that appears to suggest that what characterizes the hyperactive child's behavior is not necessarily the sheer amount of activity, but its poor organization and control.

While the supporting studies cited by Douglas (1972) are intriguing, there is still need for a more straightforward assessment of the hyperactive's impulsivity. There is a tested technique, derived from research in operant conditioning, that provides an excellent opportunity to measure the hyperactive child's ability to inhibit behavioral responding. In this procedure, termed Differential Reinforcement of Low Rate Responding (DRL), responses that occur before a set time interval has passed are not reinforced and, moreover, they serve to reset the timer governing reinforcement. Thus a subject performing on a DRL 6-second schedule must wait at least 6 seconds between responses in order to receive a reinforcement. If a response is made before the 6 seconds have elapsed, the timer will reset and another 6 seconds must pass before a successful response can be made.

Normal laboratory animals performing on this task are able to develop efficient response patterns and typically come to be reinforced for 70% to 80% of the responses they make (Hothersall, Alexander, & Slonaker, 1972). As for human DRL performance, Weisberg and Tragakis (1967) demonstrated that even young children, rating from 15 to 40 months, were able to maintain low levels of responding. Latency-age boys (Stein & Landis, 1973) and college students (Randolph, 1964) have also been shown to perform efficiently, although the samples for most studies were extremely small. One major question posed by the present study is: How would a hyperactive child fare in a DRL situation in which behavioral responding must be suppressed? Might the DRL procedure provide an objective measure of impulsivity?

Researchers investigating DRL performances have also observed that subjects in a time-based situation will often engage in a sequence or chain of behaviors between reinforceable responses. The notion is that this collateral behavior sequence helps the subject "wait out" the required temporal delay between responses. Examples of such behaviors in animals include wood chewing and tail licking (Slonaker & Hothersall, 1972). Similarly, human subjects have become convinced that these extra mediating behaviors may set up the reinforced responses (Bruner & Revusky, 1961). Experimentally precluding such collateral behaviors has led to a disruption of the subject's ability to make temporal discriminations (Stein & Landis, 1973; Stein & Flanagan, 1974). The

frequency of extra movements made by hyperactive children during a delayed reaction time task was positively related to the efficiency of their performances (Douglas, 1972; Cohen, cited in Douglas, 1972), suggesting that collateral behaviors may play a role in helping these children gain impulse control. Another intent of the current study was to examine whether or not a situation that facilitates the development of collateral behaviors helps to improve the DRL performance of hyperactive and, for that matter, normal children. And, regardless of the availability of experimentally introduced opportunities for collateral responding, will these children spontaneously elicit mediating responses?

METHOD

Subjects

Twenty hyperactive boys between 6 and 8 years of age (72-97 months) and a matched sample of nonhyperactive boys participated in the study. The subjects were selected from the client roster of a child guidance clinic; they were all nonretarded and enrolled in school at an age-appropriate grade level. A child was considered hyperactive if his teacher rated him as such on the hyperactivity factor of the Behavior Rating Scale (Conners, 1969). In accordance with the findings of a normative study (Werry, Sprague, & Cohen, 1975), those children with a total score of at least 15 on this scale were allotted to the hyperactive group, while subjects receiving a score of less than 15 made up the control group. The two groups were statistically comparable for age.

Design

The experiment essentially followed the design and methodology presented by Stein and Landis (1975). Four equal-sized groups of 10 subjects each were involved: two groups of hyperactive and two groups of nonhyperactive boys. One hyperactive and one nonhyperactive group performed on a DRL 6-second schedule without the availability of collateral buttons (DRL ONLY), while the other two groups worked on DRL using the five-button console (DRL + COLLATERAL).

Apparatus

Each subject was seated before a console containing either one (DRL ONLY) or five (DRL + COLLATERAL) microswitch buttons mounted 4.0 cm

apart. On a separate console, positioned immediately behind the button box, was a chassis containing a large six-digit add counter, which indicated the cumulative number of reinforcements (points) earned, and a red light, which lit for 2 seconds upon completion of a successful response. For the DRL ONLY condition, the one red button was programmed to produce reinforcement according to a DRL 6-second schedule. In the DRL + COLLATERAL condition, only one red button was programmed with the DRL schedule. The remaining four black collateral keys did not yield reinforcements, but they were individually monitored for the number of times they were pressed, and the response pattern was retained by an Esterline Angus Event Recorder. All sessions were videotaped in their entirety. The programming equipment, counters, timers, recorders, and monitoring equipment were housed in a separate room, which was joined to the experimental room by a one-way mirror and an intercom system.

Procedure

Prior to the testing session, the mother of every subject received an appropriate introduction to the study. The child was then taken individually to the experimental room, which contained only a table, chairs, and the testing apparatus. He then received the following instructions (adapted from Stein & Landis, 1975):

> You're going to play a game in which you will get a chance to win a lot of M&M's. Do you see this light [pointing to the reward indicator on the reinforcement box]? Every time you make this light go on you'll earn an M&M, and this counter [pointing] will keep track of how many M&M's you've won. At the end of the game I'll give you all the M&M's that you earned. Now, to make the light go on all you have to do is push this red button [pointing], and wait a little while before pressing it again. You just press this red button, wait a while, then press it again. If you press it again too soon, though, then you will not get a point, the light won't go on, and you'll have to wait a while before you can press it to get another point. But if you push the button, wait a while, then push it again, you'll get a point every time.

For the DRL + COLLATERAL groups, these instructions were added:

> These four black buttons will not give you points if you press them but, if you want, you can play with them while you're waiting to press the red button again.

The general instructions continued:

> Do you have any questions? [All questions were answered.] OK, I'll leave the room now and turn on the machine. When you hear my voice coming over the speaker saying, "You can start now," then you can begin playing the game. I'll come back when the game is over and bring you your M&M's. Have fun!

Each subject had a total of 15 minutes to earn points on the DRL schedule. At the session's end he was asked, "How did you do it?" and "Did you do

anything to help you wait?" Following his response, he was paid his due of M&M's, thanked for his participation, and returned to his mother.

Measures

The DRL task yields three measures of a subject's performance: the total number of responses, the number of reinforcements earned, and the efficiency score. Since this last score represents the percentage of reinforced responses, i.e., the number of reinforcements divided by the number of responses, it is considered to be the most sensitive indicator of a child's performance. Every subject's response, reinforcement, and efficiency scores were obtained for each of three 5-minute time blocks, and then totaled for the entire 15-minute session. Also monitored were the number of responses emitted on the collateral keys (if available) and the pattern in which they were pressed. Data on IQ, exact age, school grade, and parent ratings were also collected.

RESULTS

DRL Analysis

Analyses of variance (ANOVAs) were performed on the three operant measures, and the mean scores for each trial and for the whole session are shown in Figure 1. There was a highly significant main group (hyperactive vs. nonhyperactive) effect for all three DRL variables, especially the efficiency score, $F(1, 36) = 18.33$, $p < .0013$. Children rated as hyperactive by their teachers thus tended to emit more responses and receive fewer reinforcements during each of the time blocks than did children judged nonhyperactive. This main effect remains strong if the data are reanalyzed according to the parents' classifications of hyperactivity. The relative inefficiency of these hyperactive children suggests that they are significantly more impulsive as measured by the DRL 6-second schedule.

There were no Condition (DRL ONLY vs. DRL + COLLATERAL) effects on any of the dependent variables (efficiency score, $F(1, 36) = .133$, $p < .717$), so that whether a child played with the one-button or the five-button console had an insignificant effect upon his DRL performance. A few Group X Condition interactions did emerge: Hyperactive children in the DRL + COLLATERAL condition tended to make more responses than did their counterparts in the DRL ONLY condition, $F(1, 36) = 6.4$, $p < .016$, and nonhyperactive COLLATERAL subjects did significantly better according to the efficiency score than did the nonhyperactives in the DRL ONLY situation.

The ANOVAs yielded only one marginally significant main effect for Trials (time blocks): All groups tended to receive more rewards for their second-

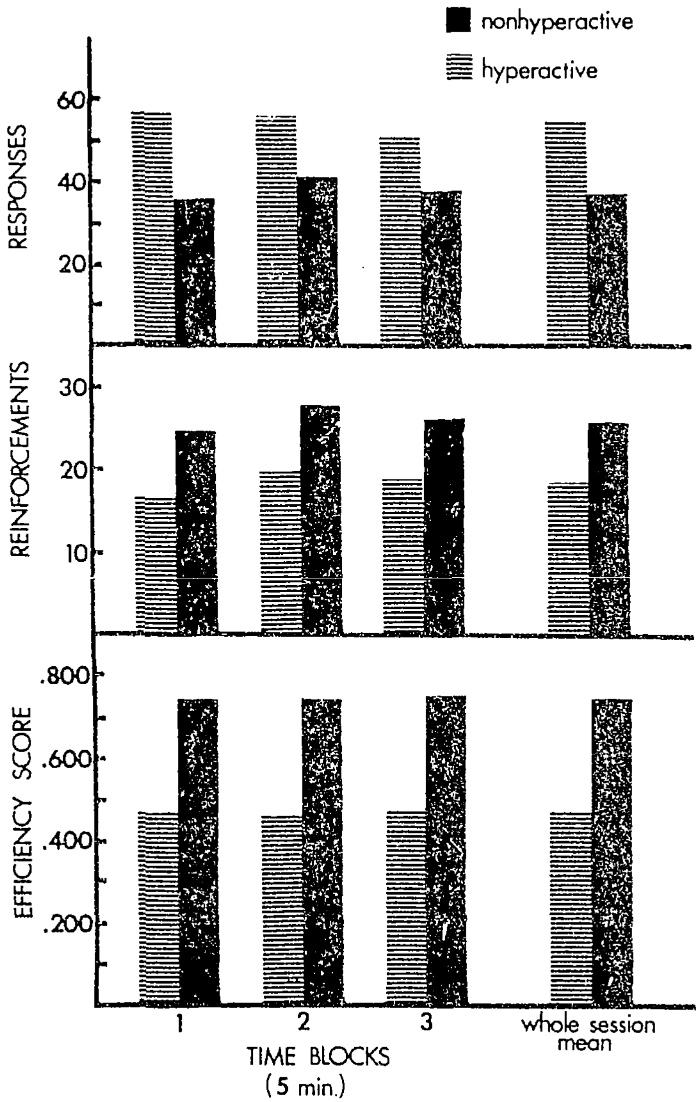


Fig. 1. The performance of hyperactive and nonhyperactive groups over three time blocks of a DRL 6-second schedule. Top: responses; middle: reinforcements; bottom: efficiency score.

trial performance, $F(2, 72) = 3.08$, $p < .051$. In general, the second trial was the most characteristic of a child's DRL work. For nearly all the various analyses, what held true for each trial taken separately, or for the session as a whole, usually held true for the second-trial measures and, again, particularly for the second time block efficiency score. This observation is supported by the intertrial correlations for each measure. Since the correlations between Trial 2 and Trial 3 tend to be quite high, mean $r = .69$, $p < .001$, it would suggest that a boy's performance was apt to stabilize by the second trial. When these interblock correlations are computed for each group separately, there is evidence that the DRL conduct of nonhyperactive subjects generally consolidated from one trial to the next as they locked into a relatively efficient pattern, mean $r = .75$, $p < .001$. The hyperactive children were more erratic across trials than the nonhyperactive children, mean $r = .54$, $p < .001$, although their performance was also relatively stable, as the intertrial correlations indicate.

In order to uncover relationships between the operant scores and a variety of descriptive data, all variables generated by this study were intercorrelated. It was first found that the subject's age is thoroughly unrelated to any other aspect of his performance or to the hyperactivity ratings supplied by the teacher or parent. The teacher ratings of hyperactivity using the Behavior Rating Scale (Conners, 1969) were significantly correlated with all DRL variables; their correlation with the whole-session efficiency score, for example, was $r = .53$, $p < .01$. The parent rating was also tied to the major DRL scores (mean $r = .35$, $p < .05$), though slightly less than were the teacher ratings. The correlation between parents and teachers, themselves, was $r = .61$ – highly significant ($p < .001$), but not so high as to instill total confidence in their reliability. Parents and teachers did closely agree (34/40 or 85% of the cases) on the gross classification of a child as hyperactive or nonhyperactive. Of the six contesting cases, parents rated four teacher-classified nonhyperactives as hyperactive and two teacher-classified hyperactives as nonhyperactive. It seems, therefore, that parents were apt to classify more children as hyperactive. Yet the teachers' mean ratings of the hyperactive boys (22.4) was slightly higher than those rendered by parents (20.0). So, teachers classified fewer children as hyperactive, but they judged the hyperactives as more severely hyperactive.

Analysis of Mediating Behaviors

As has been the case for almost all animal and human subjects working on a DRL schedule, the children who participated in this study tended to engage in a variety of interresponse behaviors that were either physically manifested or were reported to the examiner in response to a standard, end-of-session question: "How did you do it?" or "Did you do anything to help you wait?" A list of some of the mediating behaviors that were observed appears in Table

Table I. A List of Observed Mediating Behaviors

Circling DRL button with finger 9 times
Swinging legs 10, 12, or 20 times
Counting with lips
Counting out loud numbers or ABC's
Blowing on reward box
Singing out loud
Shaking reward box 10 times
Hitting knee with right hand 20 times
Foot-tapping 16 times
Tapping finger 10 times on button box
"Walking" fingers around DRL button 9 times
Stomping with foot 9 or 10 times
Running around table 1 time
Hitting side of box
Jumping jacks 4 times
Hitting collateral buttons

I. (It should be noted that only 5 of the 20 Collateral group used the collateral buttons for their interresponse regimens.) In order to be considered a mediating behavior, the sequence had to be repeated among at least 10 responses within a particular trial. Judgments concerning collateral responding were made blind from the videotapes. As it turned out, every subject engaged in, or reported, some mediating behavior. Ninety percent of the hyperactive subjects engaged in an observable mediating behavior, compared to 45% of the nonhyperactives. On the other hand, just 30% of the hyperactive boys reported the use of a nonbehavioral mediator, as opposed to 80% of the nonhyperactive subjects. A roughly comparable percentage of each group used both approaches (hyperactives = 20%, nonhyperactives = 25%).

There was an inverse correlation between the use of a physical collateral and both the efficiency ($r = -.32$, $p < .05$) and reinforcement ($r = -.55$, $p < .005$) scores, especially after the first trial. Hence, children who used an observable collateral scheme did worse on the DRL schedule; they were also apt to be rated as more hyperactive by their parents and teachers. The opposite pattern emerged for cognitive mediators as there was a very high correlation between the children's reports of counting and all measures for each trial, particularly blocks 2 and 3, mean $r = .64$, $p < .001$. Thus subjects who engaged in cognitive collaterals fared better on DRL and were usually judged less hyperactive than their noncounting counterparts. Boys who used both strategies during the session also had better scores, though not as high as the scores of those who only counted.

DISCUSSION

The prediction that children classified as hyperactive would encounter greater difficulties in managing the DRL task than would nonhyperactives was confirmed by the results of this study. According to every indicator of DRL performance, the hyperactive boys were relatively unable to refrain from making a high percentage of incorrect responses. Consequently, these results lend further credence to Douglas's (1972) proposal of an impulsivity dimension underlying the hyperkinetic syndrome, and to the validity of other measures of impulsivity presented by Sykes (cited in Douglas, 1972) and Campbell (cited in Douglas, 1972).

The experimental Condition (DRL ONLY vs. DRL + COLLATERAL) failed to emerge as a significant source of variance. However, given the observation that children engaged in many self-generated interresponse behaviors, this finding is not at all surprising. The fact is that a child will develop some type of interresponse strategy whether or not the experimenter provides a few extra buttons or telegraph keys. Indeed, in their attempts to quantify collateral responding, it appears that past investigators have missed the rich bounty of self-elicited behaviors that children exhibit during the course of a DRL session. Aside from their diversity, these schedule-induced behaviors are fascinating for the rigid manner in which they are executed; once a child chanced upon a successful strategy, he usually stuck with it ardently.

That the proclivity to engage in a physical or cognitive mediating behavior was so closely related to DRL performance as well as to ratings of hyperactivity is a highly intriguing result. Even with a relatively gross coding system, a strong pattern surfaced in which boys who displayed physical collaterals did worse on DRL and were rated more hyperactive than were boys who used cognitive means. In addition, many more hyperactive than nonhyperactive boys generated physical collaterals, an observation that may support the notion that a certain segment of the hyperactive child's hyperactivity is actually a series of mediating behaviors developed in an effort to control his impulses.

While it appears certain that cognitive mediators are more efficient strategies than physical ones, it is impossible to determine from these data whether physical collaterals actually hindered DRL efficiency or simply failed to facilitate it as much as the cognitive types did. The only hint available is the finding that the 18 subjects who employed physical collaterals during several trials had significantly higher efficiency scores than did the 9 children whose physical mediating was limited to one time block, $p < .05$. The precise nature of the interaction between a physical collateral and the DRL performance of hyperactive children is a very important avenue of research that warrants further exploration.

An unexpected aspect of these results was the extent to which the DRL variables could accurately discriminate between groups. According to the inter-trial correlations, the efficiency score can serve as a reliable and stable predictor of teacher-classified hyperactivity. Consequently, it is possible that this task could be developed into a useful clinical tool that could aid the diagnostician in the assessment of impulsivity and the classification of hyperactivity.

REFERENCES

Bruner, A., & Revusky, S. A. Collateral behavior in humans. *Journal of the Experimental Analysis of Behavior*, 1961, *4*, 349-350.

Cantwell, D. P. *The hyperactive child: Diagnosis, management, current research*. New York: Spectrum Publications, 1975.

Conners, C. K. A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry*, 1969, *126*, 152-156.

Davis, K., Sprague, R., & Werry, J. Stereotyped behavior and activity level in severe retardates: The effects of drugs. *American Journal of Mental Deficiency*, 1969, *72*, 721-727.

Doubros, S., & Daniels, G. An experimental approach to the reduction of overactive behavior. *Behaviour Research and Therapy*, 1966, *4*, 251-258.

Douglas, V. I. Stop, look and listen: The problem of sustained attention and impulse control in hyperactive and normal children. *Canadian Journal of Behavioral Science*, 1972, *4*, 259-282.

Gittis, A., & Hothersall, D. DRL performance of juvenile rats with septal lesions. *Physiological Psychology*, 1974, *2*, 38-42.

Hothersall, D., Alexander, D., & Slonaker, R. The DRL deficit of rats with septal lesions: Effects of extended training in a mediated environment. *Psychonomic Science*, 1972, *29*(1), 34-36.

Hutt, D., Jackson, P., & Level, M. Behavior parameters and drug effects: A study of a hyperkinetic epileptic child. *Epilepsia*, 1966, *7*, 250-259.

Johnson, C. F. Limits on the measurement of activity levels in children using ultrasound and photoelectric cells. *American Journal of Mental Deficiency*, 1972, *77*, 301-310.

Lee, D., & Hutt, C. A play-room designed for filming children: A note. *Journal of Child Psychology and Psychiatry*, 1964, *5*, 263-265.

McConnel, T. B., Cromwell, R. F., & Bialev, I. A. Studies in activity level: VIII. *American Journal of Mental Deficiency*, 1964, *68*, 647-651.

Ounsted, C. The hyperactive syndrome in epileptic children. *Lancet*, 1955, *269*, 303-311.

Randolph, J. J. A further examination of collateral behavior in humans. *Psychonomic Science*, 1964, *3*, 227-228.

Schulman, J. L., & Reisman, J. M. An objective measure of hyperactivity. *American Journal of Mental Deficiency*, 1959, *64*, 455-456.

Slonaker, R. L. & Hothersall, D. Collateral behaviors and the DRL deficit of rats with septal lesions. *Journal of Comparative and Physiological Psychology*, 1972, *80*, 91-96.

Stein, N., & Flanagan, S. Human DRL performance, collateral behavior and verbalization of the reinforcement contingency. *Bulletin of the Psychonomic Society*, 1974, *3*, 27-28.

Stein, N., & Landis, R. Mediating role of human collateral behavior during a spaced responding schedule of reinforcement. *Journal of Experimental Psychology*, 1973, *97*, 28-33.

Stein, N., & Landis, R. Differential reinforcement of low rates performance by impulsive and reflective children. *Journal of Experimental Child Psychology*, 1975, *19*, 37-50.

Weisberg, P., & Tragakis, C. J. Analyses of DRL behavior in young children. *Psychological Report*, 1967, *21*, 709-715.

Werry, J. S., Sprague, R. L., & Cohen, M. N. Conners' Teacher Rating Scale for use in drug studies with children: An empirical study. *Journal of Abnormal Child Psychology*, 1975, *3*(3), 217-229.

Qu K854903 84 5 11

IMAGE EVALUATION TEST TARGET (MT-3)

150mm

6″

PHOTOGRAPHIC SCIENCES CORPORATION
770 BASKET ROAD
P.O. BOX 338
WEBSTER, NEW YORK 14580
(716) 265-1600

K854903

Journal of Abnormal Child Psychology, Vol. 12, No. 4, 1984, pp. 561-572

Performance of Disturbed Hyperactive and Nonhyperactive Children on an Objective Measure of Hyperactivity[1]

F. Daniel McClure
David C. Wilson Neuropsychiatric Hospital

Michael Gordon[2]
SUNY, Upstate Medical Center

Twenty hyperactive emotionally disturbed children (6-11 years) and a matched sample of nonhyperactive emotionally disturbed children were selected from the population of a therapeutic day treatment facility on the basis of teacher ratings. They were administered the Matching Familiar Figures Test-20 and were rated on several scales of impulsivity and/or hyperactivity. Each subject was required to perform on the Delay Task of the Gordon Diagnostic System, which required them to inhibit behavioral responding on a temporally based schedule (DRL-6) in order to win points. Children classified as hyperactive, whether by one or more criteria, were relatively unable to refrain from emitting a high number of nonreinforced responses. Moreover, these performance differences persisted regardless of age or IQ and were stable over the 8 minutes required to complete the test.

Manuscript received in final form January 4, 1984.
[1]The authors gratefully acknowledge the efforts of the following members of the Institute of Clinical Psychology, University of Virginia: Richard Abidin, Ann Loper, Ronald Reeve, and Sherry Kraft. Data analyses were conducted using STATMANAGER, Hayden Software Company, copyright 1983.
[2]Address all correspondence to Michael Gordon, Upstate Medical Center, 750 East Adams Street, Syracuse, New York 13210.

0091-0627 84 1200-0561$03.50 0

While current efforts to understand hyperactive behavior in children have highlighted deficits in attentional processes (Douglas & Peters, 1979), there has been mounting evidence that problems with impulse inhibition are also central to the disorder. In her review of the literature, Douglas (1972) concluded that hyperactive children "are apparently unable to keep their impulses under control in order to cope with situations in which care, concentration, or organized planning are required" (p. 275). Impulsive errors have been found to occur at substantially higher rates among hyperactive children on continuous performance tasks (Sykes, 1969; Sykes, Douglas, Weiss, & Minde, 1971; Sykes, Douglas, & Morganstern, 1973) as well as on delayed reaction time tasks (Cohen, 1970; Cohen, Douglas, Morganstern, 1972; Firestone & Douglas, 1975; Parry & Douglas, 1983), along with errors suggestive of lapses in attention. This dimension of impulsivity, which Douglas (1972) refers to as the "stop, look, and listen" dimension, is pervasive among these children and profoundly colors their behavior.

Measures of impulsivity in children currently available fall into two basic categories, those that are largely limited to laboratory investigations and those that were originally intended for the assessment of other behavioral domains but that have been adapted to assess impulsivity. Many clinicians rely upon adaptations of the Bender Gestalt test (Koppitz, 1964; Tolor & Branningan, 1980) and selected subtests of the WISC-R (Kaufman, 1979) to identify and assess impulsivity. The validity of these measures suffers because they assess a wide variety of rather loosely defined cognitive abilities that may or may not be affected by the dimension of impulsivity (Douglas, 1972).

Instruments such as the Continuous Performance Task (Rosvold, Mirsky, Sarason, Bransome, & Beck, 1956) and delayed reaction time tasks (Cohen, 1970; Cohen et al., 1972; Firestone & Douglas, 1975) have demonstrated a high degree of discriminative ability with hyperactive children, but have been largely limited to use in the laboratory. Rating scales such as the Teacher Rating Scale (Conners, 1969) have enjoyed widespread use in both research and clinical application with hyperactive children despite their inherent difficulties with rater bias, the "halo" effect, and the relative poor differentiation of the quantitative ratings involved (Ross & Ross, 1976).

Gordon (1979) has developed a methodology stemming from research in operant conditioning that has shown considerable promise as a stable and reliable measure of impulsive styles. This technique uses the Differential Reinforcement of Low Rate Responding (DRL) first described by Skinner (1938). The DRL schedule provides for reinforcement of a response that occurs after a certain specified time has elapsed since the emission of the previous response. Moreover, responses that occur prior to the termination of this time interval are not reinforced and serve to reset the timer governing the time interval. Thus, a subject performing on a DRL 6-second schedule would receive a reinforcement for every response emitted after an interval of 6 seconds, but responses emitted within 6 seconds of the previous response would not be reinforced, and the subject would have to wait another 6 seconds before the next response emitted could be reinforced.

In a study of 20 hyperactive boys and a matched sample of nonhyperactives, designated as such on the basis of the Teacher Rating Scale (Conners, 1969), Gordon (1979) found that this technique differentiated the two groups at a high level of statistical significance. The most sensitive measure was the efficiency ratio (ER), which is derived by dividing the total number of correct responses by the total number of responses. In addition, the nature of self-generated mediational strategies employed by each subject in his adaptation to the DRL schedule provided valuable data. Behaviors were rated as covert, or cognitive (e.g., counting silently), on the basis of posttest inquiry, or as overt, or behavioral, based on observer ratings. Statistical analysis of these collateral behaviors indicated that subjects who developed more cognitively oriented mediational strategies tended to achieve a significantly more efficient adaptation to the DRL schedule than did those subjects whose attempts at mediating were more behaviorally manifest.

An outgrowth of this research was the development of the Gordon Diagnostic System (GDS), which includes, in part, a version of the original DRL paradigm. (It also administers a version of the Continuous Performance Test.) In contrast to the electromechanical apparatus originally utilized, the GDS is a portable device that performs the DRL task and records the data in a single unit. This study was an attempt to determine if any performance differences would emerge in the use of this solid-state version of the task, and to extend Gordon's original study to include children with a wide variety of psychosocial, emotional, and psychological disturbances.

METHOD

Subject

Twenty emotionally disturbed children (18 males and 2 females) classified as hyperactive on the basis of the Teacher Rating Scale (TRS; Conners, 1969) and a matched sample of children rated as nonhyperactive (16 males and 4 females) were selected from the roster of a therapeutic day treatment facility.

Following from a normative study (Werry, Sprague, & Cohen, 1975), the criterion for designation as hyperactive was a mean rating of greater than 1.5 on the TRS. The two groups were statistically comparable for age, IQ, and SES (See Table I). While there was no method of controlling for severity of disturbances, the primary diagnoses for all children in this facility were conduct and emotional disorders, both in conjunction with cognitive and learning deficits.

Apparatus

In addition to the TRS, ratings were obtained from two other sources, the Self-Control Rating Scales (SCRS; Kendall & Wilcox, 1979) and the Teacher Report Form (TRF; Achenbach & Edelbrock 1983). The SCRS is a 33-item scale of impulsive behaviors, rated by teachers on a 7-point continuum. The TRF is a 113-item scale designed to assess a wide variety of behavioral problems and competencies. Behaviors are rated on a 3-point scale, with a rating of 0 representing a behavior that is "not true" of the child in contrast to a 2-point description of "very true."

Each child was administered the Matching Familiar Figures Test-20 (Cairns & Cammock, 1978), which is a longer, more reliable version of the original Matching Familiar Figures Test (Kagan, 1966). Each child also was administered the Bender Gestalt Test (Bender, 1938), which was scored for impulsivity (Tolor & Brannigan, 1980).

The Gordon Diagnostic System (GDS) is a one-button solid-state console manufactured by Clinical Diagnostics, Inc. It is designed to provide a behavioral measure of the child's ability to inhibit responding. The 6-second delay interval was employed in the current study. The GDS yields three fundamental sets of data: the absolute number of responses, the total number of rewarded responses, and an efficiency ratio (ER), which is obtained by dividing the total number of rewarded responses by the absolute number of responses. The ER thus represents the percentage of correct responses. If the child makes an unusally low or high number of responses, the ER may be a somewhat misleading index of performance. These three scores are computed for each 2-minute time block and then totaled for the overall 8-minute trial. During GDS performance each child's behavior was also recorded by the test administrator to assess the presence or persistence of any overt mediational strategies.

Procedure

Teachers initially completed the TRS and then, after a period of 4 weeks, the SCRS. The time interval was to minimize any possible overlap

Table I. Means and Standard Deviations for Hyperactive and Nonhyperactive Groups for Age, IQ, and TRS Factors

	Age in months	IQ WISC-R	Factor I	Factor II	Factor III	Factor IV
Hyperactive (n = 20)						
Mean	107.4	86.7	1.33	1.58	1.01	2.36
SD	26.2	11.85	.54	.60	.48	.42
Nonhyperactive (n = 20)						
Mean	102.9	82.9	.80	1.19	.93	1.19
SD	23.3	11.1	.48	.57	.43	.38

between ratings on the two scales. Head child care workers completed the TRF. The order of administration of the MFF-20 and the GDS was reversed for each group in order to control for the effects of boredom and/or fatigue. The Bender was always given first in the battery. The examiner was blind to each subject's group status.

Each subject was taken into the testing room individually and seated at a table. The MFF-20 was administered in standard fashion, and the GDS was introduced as a game, much like computerized arcade games. Prior to the GDS administration each subject was given the following instructions (adapted from Gordon 1979).

> Your are going to play a game in which you will get a chance to win a lot of points. Do you see this red light [pointing]? Every time you make this red light go on you'll earn a point and this counter [pointing] will keep track of how many points you've earned. Now, to make this light go on, all you have to do is push this blue button [pointing] and wait a while before pressing it again. You just push the blue button, wait a while, then press it again. If you press it too soon, though, you will not get a point, the red light won't go on, and you'll have to wait a while before you can press it to get another point. But, if you press the button, wait a while, and press it again, you'll get a point every time.

The examiner then answered any questions and made sure that the child had understood the instructions. Moving to a position adjacent to and somewhat behind the child the examiner continued:

> I'll sit here until you're done. You'll know when the game is over because this light will come on again. If you have any questions while you're playing the game please wait to ask them until the game is over. We can talk about the game when you're finished.

At the end of the trial each child was asked two questions: "Did you enjoy the game?" and "What did you do to help you wait?" The former was designed to address issues of motivation to perform well, while the latter was designed to address the issue of nonobservable mediational strategies.

RESULTS

To test the hypothesis that the GDS Delay Task could differentiate between the two groups, a one-tailed t test was conducted for each GDS measure and group membership (hyperactive vs. nonhyperative). The ER appeared to be the most sensitive measure, $t(38) = 5.67, p < .001$, and the absolute number of responses emitted also differentiated the two groups at a convincing level $t(38) = 4.26, p < .01$. Only the total number of rewarded responses failed to differentiate the two groups at a level of statistical significance, $t(38) = 1.79, p < .09$. While the two groups did not differ significantly on the number of times they were rewarded, those rated as hyperactive emitted a significantly greater number of nonrewarded responses. The mean ER for the nonhyperactive group was .68 (SD = .15), while the mean ER for the hyperactive group was .42 (SD = .14). Since the ER is clearly derived from the other two scores, no comparisons of dependent variables were conducted.

To further judge the ability of the GDS to differentiate between children rated as hyperactive or nonhyperactive, post hoc analyses were performed with groups formed on the basis of being identified as hyperactive or nonhyperactive on four criterion measures (TRS, SCRS, MFF-20, and TRF). The Bender was not included as a criterion measure because the drawings of many of these subjects were so disturbed as to render them unscorable. The criterion for hyperactivity on the MFF-20 was a mean latency to first response of less than 10 seconds. Criteria for the TRF and SCRS were determined by taking the mean SCRS raw score and mean T score for the TRF hyperactivity factor for the entire sample, and then excluding all subjects within one standard deviation of the respective means. For subjects who were identified as hyperactive (n = 8) or nonhyperactive (n = 8) on the basis of all four criterion measures, there was agreement between the GDS and criterion measures in 15 cases (84%, Fisher's Exact Test, p = .01). The GDS criterion score was less than .54. When groups were constructed by subjects classified as hyperactive or nonhyperactive by *at least* three of the four criterion measures, there was agreement in 29 of 32 cases (91%, Fisher's Exact Test, p = .001).

Correlations presented in Table II depict the relationships between the three GDS variables and all the other measures. Although the criterion measures are related to one another to the extent of statistical significance or near statistical significance, only Factor IV of the TRS and both MFF-20 variables are significantly related to GDS performance.

Correlations between age and IQ for three GDS variables across each of the 2-minute time blocks as well as for the 8-minute trial yielded no statistically significant relationships. Correlations were also obtained for each GDS variable across each of the 2-minute blocks and for the total trial in order to assess the internal stability of the measure across time. While subjects rated as hyperactive demonstrated somewhat more variability in their performance across the 8 minutes, the patterns of responding were quite stable for both groups. Another method employed to look at the stability of the instrument in terms of discriminative ability was to run multiple t tests between the two groups for each GDS variable in each time block. Again, the absolute number of responses emitted and the ER consistently differentiated between the two groups, regardless of what particular discrete time block was used for comparison.

t tests conducted between TRS factors for both groups indicate that while Hyperactivity (Factor IV) was by far the most powerful in terms of differentiating the two groups, $t(38) = 9.28, p < .001$, the two groups also differed significantly along the dimensions of Conduct Problem (Factor I),

Table II. Correlational Matrices for the Three GDS Variables with All Other Measures

	Responses	Correct	ER	MFF-L	MFF-E	TRS-I	TRS-II	TRS-III	TR-SIV	TRF	SCRS
Responses											
Total correct	−.31										
ER	−.82[b]	.62[b]									
MFF-L	−.37[a]	.22	.46[b]								
MFF-E	.51[b]	−.11	−.67[a]	−.46[b]							
TRS-I	.16	−.20	−.28	−.01	.25						
TRS-II	−.04	.08	.02	−.09	−.12	−.19					
TRS-III	−.35[a]	.16	.27	.05	−.06	.22	.47[b]				
TRS-IV	.57[b]	−.21	−.59[b]	−.39[b]	.23	.43[b]	.36	.16			
TRF	.17	−.05	−.26	−.34	.17	.31	.06	.09	.48[b]		
SCRS	.02	.08	−.16	−.27	.03	.19	.13	.26	.42[b]	.56[b]	

[a] $p < .05$.
[b] $p < .01$.

$t(38) = 3.28, p > .001$, and Inattentiveness (Factor II), $t(38) = 2.11, p < .05$. This pattern of TRS scores is highly consistent with the pattern reported by Werry et al. (1975). Thus, it appears that although hyperactivity is the primary dimension measured by the TRS, it is also sensitive to related disturbances in attention and conduct. However, of these four factors, only Factor IV was significantly correlated with the GDS Efficiency Ratio, $r = .58, p < .01$.

Statistical analyses of mediational strategies were not conducted due to a paucity of data. Very few subjects engaged in any particular behavior(s) consistently enough for them to be considered mediational in nature, and few subjects were able to verbalize strategies that they may have employed.

DISCUSSION

The results lend considerable support to the GDS as a stable and reliable method of differentiating teacher-rated hyperactive from nonhyperactive children. Whether classified by the TRS alone or in conjunction with the other criteria, children classified as hyperactive were relatively unable to perform efficiently on this task. Although their average number of rewarded responses were not significantly different, the hyperactives were unable to refrain from emitting a high number of nonrewarded responses. Since there is a strong need for reliable behavioral measures in the overall assessment of hyperactivity (Prout & Ingram, 1982), the GDS may represent an important step forward.

Despite substantial qualitative differences in samples, the performances of the subjects in the present study very closely approximated those in Gordon's (1979) original study; both groups tended to obtain ERs lower than those reported earlier. Thus, it may be that both samples in this study actually consisted of hyperactive children who differed only in their degree of hyperactivity. However, if this is the case, the discriminative ability of the GDS is only enhanced, as it differentiated the two groups within a much more restricted range than would otherwise be the case. In the earlier study, the number of rewarded responses also differentiated the two groups at a statistically significant level. Here both groups also tended to make fewer rewarded responses. This difference may also relate to qualitative differences in samples between the two studies. Nevertheless, subjects rated as hyperactive in both studies demonstrated a significantly less efficient adaptation to the task than did their nonhyperactive counterparts.

The similarities in the performance of the subjects in this study and those in the original study persist despite some very important differences in subject selection and methodology. First, the miniaturization of the apparatus did not seem to affect the discriminative ability of the DRL procedure. Another important difference has to do with the nature of reinforcement. In the original study, each subject received an M & M for each rewarded response, and these were dispensed at the end of the session. In the present study, there were no tangible reinforcers.

Finally, this study was conducted with children whose emotional and behavioral difficulties are of such a magnitude as to require very restrictive education/therapeutic placement. Within this sample, the two groups differed significantly on three of the TRS factors—Conduct Disorder, Inattentive-Passive, and Hyperactivity. The relationships among these factors are consistent with the findings of several studies (Sprague et al., 1974) and would seem to support the notion that underlying deficits in attention and impulse inhibition contribute to behaviors that bring these youngsters into conflict with their environments. The GDS appears to provide information about the extent to which a child, regardless of age, IQ, or specific diagnostic classification, displays such an "impulsive style."

The fact that, in this sample, age and IQ were unrelated in any way to Delay Task performance has very clear methodological and conceptual implications. A chronic difficulty with other measures of impulsivity, such as the Matching Familiar Figures Test (Cairns & Cammock, 1978), is that they are closely related to intellectual functioning and developmental status. Consequently, lack of consistency between measures of impulsivity and hyperactivity reported in some studies (Sandberg et al., 1978) may be related more to error of measurement than to the absence of a dimension of underlying impulsivity per se. It may also reflect that impulsivity is not a unitary construct (Paulsen & Johnson, 1980) and that the various tasks assess different aspects of a child's functioning as it relates to impulsive behavior. Children may be impulsive in different ways and to different extents as a result of variations in environmental, cognitive, and/or constitutional factors. In some children these deficits are expressed primarily through motoric channels while, in others, the expression may be more cognitive in nature.

It should also be pointed out that, although the Delay Task would seem to tap most broadly into the area of impulsivity, there are certainly a myriad of processes at play while the child performs the test. These most likely include motivational factors and time estimation ability, as well as the capacity to develop an efficient strategy. Further research into the interplay of these processes is in progress.

REFERENCES

Achenbach, T. M. (1978). The child behavior profile. I. Boys aged 6-11. *Journal of Consulting and Clinical Psychology, 46*, 478-488.

Achenbach, T. M., & Edelbrock, C. (1983). *Mannual for the Child Behavior Checklist and Revised Child Behavior Profile.* New York: Queen City Printer.

Bender, L. (1938). A visual motor Gestalt test and its clinical. *American Orthopsychiatry Association Research Monographs, 3.*

Carins, E., & Cammock, T. (1978). Development of a more reliable version of the Matching Familiar Figures Test. *Developmental Psychology, 14,* 555-560.

Cohen, N. J. (1970). *Psychophysiological concomitants of attention in hyperactive children.* Unpublished manuscript, McGill University.

Cohen, N. J., Douglas, V. I., Morganstern, G. (1972). The effect of methylphenidate on attentive behavior and autonomic activity in hyperactive children. *Psychopharmacologia, 22,* 282-294.

Conners, C. K. (1969). A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry, 126,* 884-888.

Douglas, V. I. (1972). Stop, look, and listen: The problem of sustained attention and impulse control in hyperactive children. *Canadian Journal of Behavioral Sciences, 4,* 259-282.

Douglas, V. I., & Peters, K. G. (1979). Toward a clearer definition of attentional déficits of hyperactive children. In G. A. Hale & M. Lewis (Eds.), *Attention and the development of cognitive skills* (pp. 173-247). New York, Plenum.

Firestone, P., & Douglas, V. I. (1975). The effects of reward and punishment on reaction times and autonomic activity in hyperactive and normal children. *Journal of Abnormal Child Psychology, 3,* 201-215.

Gordon, M. (1979). The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive boys. *Journal of Abnormal Psychology, 7,* 317-326.

Kagan, J. (1966). Reflectivity-impulsivity: The generality and dynamics of cognitive tempo. *Journal of Abnormal Psychology, 71,* 17-24.

Kaufman, A. S. (1979). *Intelligent Testing with the WISC-R.* New York: Wiley.

Kendall, P. C., & Wilcox, L. E. (1979). Self-control in children: Development of a rating scale *Journal of Consulting and Clinical Psychology, 7,* 317-326.

Kopitz, E. M. (1964). *The Bender Gestalt Test for Young Children.* New York: Grune and Stratton.

Parry, P. A., & Douglas, V. I. (1983). Effects of reinforcement on concept identification in hyperactive children. *Journal of Abnormal Child Psychology, 11,* 327-340.

Paulsen, K., & Johnson, M. (1980). Impulsivity: A multidimensional concept with developmental aspects. *Journal of Abnormal Child Psychology, 8,* 269-277.

Prout, H., & Ingram, R. E. (1982). Guidelines for the behavioral assessment of hyperactivity. *Journal of Learning Disabilities, 15,* 393-395.

Ross, D. M., & Ross, S. A. (1976). *Hyperactivity: Research, theory, action.* New York: Wiley.

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, E., & Beck, L. A. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology, 20,* 343-352.

Sandberg, S. T., Rutter, M., & Taylor, E. (1978). Hyperkinetic disorder in psychiatric clinic attenders. *Developomental Medical and Child Neurology, 20,* 279-299.

Skinner, B. F. (1938). *The behavior of organisms.* New York: Appleton-Century-Crofts.

Sykes, D. H. (1969). *Sustained attention in hyperactive children.* Unpublished doctoral dissertation, McGill University.

Sykes, D. H., Douglas, V. I., & Morganstern, G. (1973). Sustained attention in hyperactive chilchildren. *Journal of Child Psychology and Psychiatry, 14,* 213-220.

Sykes, D. H., Douglas., V. I., Weiss, G., & Minde, K. K. (1971). Attention in hyperative children and the effect of methylphenidate (Ritalin). *Journal of Child Psychology and Psychiatry, 12,* 129-139.

Tolor, A., & Brannigan, G. (1980). *Applications of the Bender-Gestalt Test.* Springfield, Illinois: Thomas.

Werry, J. S., Sprague, R. L., & Cohen, M. N. (1975). Conners' teacher rating scale for use in drug studies with children—An empirical study. *Journal of Abnormal Child Psychology, 3,* 217-229.

ASSESSMENT OF ATTENTION DEFICIT DISORDERS USING THE GORDON DIAGNOSTIC SYSTEM

Michael Gordon, Ph.D.
State University of New York, Upstate Medical Center

F. Daniel McClure, Ph.D.
David C. Wilson Neuropsychiatric Hospital

Presented at the 92nd Annual Convention of the American Psychological Association at Toronto, Ontario, Canada, August 1984

Correspondence address:
750 East Adams Street
Syracuse, New York 13210
(315) 473-8145

ASSESSMENT OF ATTENTION DEFICIT DISORDERS USING THE GORDON DIAGNOSTIC SYSTEM

Abstract

Despite large numbers of referrals for evaluations of Attention Deficit Disorders/Hyperactivity, school psychologists have had to contend with assessments conducted without the aid of objective and reliable procedures. The Gordon Diagnostic System (GDS) administers tasks which provide data on a child's ability to delay and to sustain attention (Gordon, 1979; Gordon & McClure, 1983; Gordon & McClure, 1984; McClure & Gordon, 1984). We have now gathered normative data on 700 children between 3 and 16 years of age. Studies are reported which compare diagnoses made by traditional approaches with those based on GDS performance. An analysis of the time expended in traditional assessment activities is also presented. In addition, we describe a study contrasting the GDS performances of children classified as severely learning disabled, educationally mentally handicapped, and emotionally handicapped. Finally, we introduce initial normative data on a new GDS task which provides information about the effects of distraction on sustained attention. The cost effectiveness of traditional evaluations as well as implications for future diagnostic evaluations are explored.

ASSESSMENT OF ATTENTION DEFICIT DISORDERS USING THE GORDON DIAGNOSTIC SYSTEM

INTRODUCTION

According to a recent survey (Gordon & McClure, 1983), approximately 30% of referrals to school psychologists for psychological testing contain a question of hyperactivity. Of these referrals, practioners reported that only about 17% were eventually diagnosed as "truly" hyperactive. It was documented that the school psychologists relied far more on clinical judgment than they did on data from psychological tests which they considered relatively irrelevant to the diagnosis. Comments by the respondents reflected a general lack of confidence in the evaluation of this disorder.

The Gordon Diagnostic System (GDS) was introduced to provide practitioners with a means of obtaining accurate data on a child's ability to delay and to sustain attention. It has been shown that the GDS accurately discriminates between groups classified as hyperactive and nonhyperactive in samples from an outpatient clinic (Gordon, 1979), a day treatment center for severely emotionally disturbed children (McClure & Gordon, 1984), and a school-referred population (Gordon & McClure, in press). It also has been shown that the GDS Delay Task is sensitive to the effects of stimulant medication (Shue & Douglas, 1983).

DESCRIPTION OF THE GDS

The GDS is a portable, electronic unit which allows for the administration of two tasks. The Delay Task requires the child to inhibit responding in order to gain a point. Specifically, the child is instructed to press a button, wait, and then press the button again. If he/she refrains from responding for the delay interval, a light flashes and a reward counter increments. If the child responds before the interval elapses then the timer resets and no reward points are recorded. For children 6 years of age and older, a delay interval of 6 seconds is used and the session lasts 8 minutes. Preschoolers are required to wait 4 seconds across a 6-minute session.

The GDS Delay Task generates three raw scores: the number of responses (button presses), the number of correct responses (i.e. Rewards) and the Efficiency Ratio which represents the percentage of correct responses.

The GDS also contains the Vigilance Task which is a version of the Continuous Performance Test (Rosvold, et al., 1956). The child is presented with a series of digits on the front display and is told to press the button every time a "9" appears that had been preceded by a "1". The GDS records the number of correct presses, the number of times the child fails to press the button when the "1/9" combination appears (i.e. errors of omission), and the number of extraneous button presses (i.e. errors of commission). For the testing of younger children, the GDS contains a "1" mode, which requires the subject to press the button every time a "1" appears. The same performance measures are recorded. The Vigilance Task lasts 9 minutes for children 6 years and older, and 6 minutes for preschoolers.

The GDS Distractibility Task is a new procedure we are developing to assess the effects of distraction on a child's ability to sustain attention. It is identical to the Vigilance Task except that random numbers appear at random intervals on each side of the "hot" digit. The same set of scores are recorded for this task as for the Vigilance Task.

A microcomputer software package is also available that automatically receives data from the GDS, tabulates the various scores, stores the information, and prints out an interpretive summary of the performance. Included in the report are educational recommendations and suggestions for pharmacotherapy.

THE GDS IN THE ELEMENTARY SCHOOL -- Current Studies

Prior research (Gordon, 1979; Gordon & McClure; 1983,1984; McClure & Gordon, 1984) established preliminary normative data and demonstrated the diagnostic accuracy of the GDS. We now have greatly expanded the standardization sample and have gathered data on clinical use of the GDS in actual school settings.

A. The Standardization Study -- Normative data for the Delay and Vigilance Tasks are based upon the testing of 700 children while pilot data for the Distractibility Task come from the performances of 58 children. Subjects were selected at random from nine public and private schools in New York State and Virginia. To be included in the sample the child had to be rated by the teacher as nonhyperactive on the Teacher Rating Form (Edelbrock & Achenbach, 1984) or Teacher Rating Scale (Conners, 1969), and have no history of grade retention. Data on IQ, achievement, and SES were also collected.

For ease of interpretation, we have arranged the standardization data into Threshold Tables (see Tables 1-3). Based upon data from standardization projects, we have divided scores into three ranges. A score is considered ABNORMAL if it is typical of less than 5 percent of the normal population, i.e. the 5th percentile or less. Scores in the BORDERLINE range are those which fall between the 6th and 25th percentile. Finally, scores that are above the 25th percentile are classified as NORMAL. These cutoffs correspond closely to clinical experience and our research studies in that children classified as hyperactive by other measures typically perform in the ABNORMAL range of GDS scores.

B. The Cost-Effectiveness and Accuracy of Traditional Evaluations for ADD/Hyperacitvity -- Prior studies have shown that the GDS classifies children as hyperactive or nonhyperactive with a high level of accuracy when behavior rating scales and research tasks are used as criteria (McCLure & Gordon, 1984). At the same time, numerous studies have demonstrated that traditional assessment techniques such as the WISC-R and the Bender-Gestalt Designs, do a poor job at discriminating between groups classified as hyperactive and nonhyperactive (Sandoval, 1977). The pitfalls of relying on clinical judgment have also been documented (Hughes, Goldman, & Snyder, 1983; Marquis, 1983).

To explore the ramifications of current practices in the evaluation of ADD/Hyperactivity, we asked 3 highly-experienced school psychologists to record the amount of time they spent conducting assessments of 31 children referred with hyperactivity as a presenting complaint. The instruments administered during the course of their assessments included the Wechsler Intelligence Scale for Children-Revised (WISC-R), Bender Gestalt Designs, Children's Personality Questionnaire, Peabody Individual Achievement Test, Visual Aural Digit Span Test, Sentence Completion, House-Tree-Person, and Draw-A-Person. The results, presented in Table 4, indicate that each evaluation occupied between 10 and 16 hours of the psychologist's time.

Although their time was distributed across several categories of activity, they spent the most number of hours actually testing the child.

After the psychologists had arrived at a diagnosis based on their assessment, each child was tested with the GDS. The children were classified as ADD by the GDS criteria if their scores fell within the ABNORMAL range on one or both of the GDS Tasks. The agreement between classifications conducted with and without the GDS is illustrated in Table 5. Using the GDS scores as criteria, the results indicate that evaluations by traditional assessment tended to seriously overdiagnose ADD. The full assessment battery did not appear to miss children who were truly ADD but that finding may have been affected by the nature of the sample. Upon further inquiry after completion of the study, it appeared that the children considered by the GDS scores to be misdiagnosed actually represented those cases about which the clinicians were least confident in their classifications.

C. The GDS Performance of Children in Self-Contained Classrooms -- The purpose of this study was to determine the prevalence of attentional disorders in a sample of children placed in self-contained classrooms. The data presented here actually represent the initial findings of a large study of GDS performance in school-based populations of retarded, blind, deaf and emotionally disturbed children. The children in the current sample had been labelled by the

Committee on the Handicapped as either Severely Learning Disabled (SLD; n=13), Educationally Mentally Handicapped (EMH; n=7), or Emotionally Handicapped (EH; n=9). The results (Table 6) indicated that nearly a quarter of the sample scored abnormally on both Tasks and would be considered significantly impulsive and inattentive. The highest percentage of abnormal GDS performances came from the EH group where nearly 55% had one or both scores in the ABNORMAL range. Inspection of teacher rating scales and discussions with the teachers themselves indicated that approximately 75% of this group had overactivity as a presenting complaint. The teachers also suggested that information about a particular child's level of impulsivity and inattentiveness helped them develop more systematic remediation programs.

Summary

Our early investigations with the GDS centered mostly around clinic populations, and the preliminary norms were based on a relatively small sample. The normative data are now quite extensive and we have had considerable experience with using the GDS in school settings. The findings presented here indicate that school psychologists, using traditional assessment techniques, expend a great deal of time evaluating children for ADD/Hyperactivity. There is also the suggestion that classifications based on traditional assessment lead to an overdiagnosis of the problem. This would seem to indicate the possibility that many children classified as ADD/hyperactive either did not require treatment or might have benefitted from other forms of remediation. Finally, we have found that children in self-contained classrooms can be evaluated with the GDS. A significant percentage of the sample showed signs of impulsivity and inattentiveness.

ACKNOWLEDGMENTS

These studies were possible because of the cooperation of the Liverpool School District, Liverpool, NY. Thanks are due particularly to Mr. Kevin Bacher, Mr. William Mugridge, Ms. Susan Venditto, Ms. Gerry Mouio, and Mr. William Cleary.

We are greatly indebted to Ms. Barbara Mettelman, Research Coordinator for GDS Research Projects, Dr. Ernest M. Post, Dr. Wendy Gordon, Mr. Jeffrey Weiner, Ms. Sandy Tavitian, and Ms. Lacie DeHimer for their active involvement in our research program.

This research was funded by a grant to the Research Foundation of the State University of New York awarded by Clinical Diagnostics, Inc.

NOTE:

The GDS is commercially available from Clinical Diagnostics, Inc., 1536 Cole Blvd. Suite 350, Golden, Colorado 80401 (Phone: 800-521-4503).

References

Conners, C.K. (1969). A teacher rating scale for use in drug studies with children. American Journal of Psychiatry, 126, 884-888.

Edelbrock,C. & Achenbach, T.A. (1984). The Teacher Version of the Child Behavior Profile: I. Boys aged 6-11. Journal of Consulting and Clinical Psychology, 52, (2), 207-217.

Gordon, M. (1979). The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive boys. Journal of Abnormal Child Psychology, 7 (3), 317-326.

Gordon, M. & McClure, F.D. (1983, August). The objective assessment of attention deficit disorders. Paper presented at the 91st Annual Convention of the American Psychological Association, Anaheim, California.

Gordon, M. & McClure, F.D. (1983). Survey of current practices and problems in the assessment of attention deficit disorders. Manuscript submitted for publication.

Hughes, M.C., Goldman, B.L., Snyder, N.F. (1983). Hyperactivity and the attention deficit disorder. American Family Physician, 27, 119-126.

Marquis, P. (1983). The attention deficit disorder. Postgraduate Medicine, 73, 295-300.

McClure, F.D. & Gordon, M. (1984). Performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. Journal of Abnormal Child Psychology, 12, (4), 561-571.

Rosvold, H.E., Mirsky, A.F., Saranson, I., Bransome, E. & Beck, L.A. (1956). A continuous performance of brain damage. Journal of Clinical Psychology, 20, 343-352.

Sandoval, J. (1977). The measurement of the hyperactive syndrome in children. Review of Educational Research, 47, 293-318.

Shue, K. & Douglas, V.I. (1983). [The effect of Ritalin on hyperactive children's performance on a DRL measure of impulsivity]. Unpublished raw data.

Table 1. Threshold Table for the Delay Task

3-5 YEAR OLDS (n=138)

	TOTAL ER	TOTAL REWARDS	TOTAL RESPONSES
NORMAL	.61 - 1.0	≥ 46	≤ 64
BORDERLINE	.39 - .60	16 - 25	65 - 91
ABNORMAL	≤.38	≤ 15	≥ 92 & ≤ 20

6-7 YEAR OLDS (n=139)

	TOTAL ER	TOTAL REWARDS	TOTAL RESPONSES
NORMAL	.73 - 1.0	≥ 35	≤ 62
BORDERLINE	.42 - .72	14 - 34	63 - 75
ABNORMAL	≤.41	≤ 13	≥ 76 & ≤ 20

8-11 YEAR OLDS (n=269)

	TOTAL ER	TOTAL REWARDS	TOTAL RESPONSES
NORMAL	.79 - 1.0	≥ 43	≤ 66
BORDERLINE	.55 - .78	31 - 42	67 - 76
ABNORMAL	≤.54	≤ 30	≥ 77 & ≤ 20

12-16 YEAR OLDS (n=154)

	TOTAL ER	TOTAL REWARDS	TOTAL RESPONSES
NORMAL	.86 - 1.0	≥ 47	≤ 66
BORDERLINE	.71 - .85	39 - 46	67 - 75
ABNORMAL	≤.70	≤ 38	≥ 76 & ≤ 20

Table 2. Threshold Tables for the Vigilance Task

3-5 YEAR OLDS (n=138)

	TOTAL CORRECT	TOTAL COMMISSIONS
NORMAL	21 - 29	0 - 5
BORDERLINE	13 - 20	6 - 24
ABNORMAL	≤ 12	≥ 25

6-7 YEAR OLDS (n=137)

	TOTAL CORRECT	TOTAL COMMISSIONS
NORMAL	33 - 45	≤ 10
BORDERLINE	22 - 32	11 - 23
ABNORMAL	≤ 21	≥ 24

8-9 YEAR OLDS (n=136)

	TOTAL CORRECT	TOTAL COMMISSIONS
NORMAL	40 - 45	≤ 8
BORDERLINE	33 - 39	9 - 21
ABNORMAL	≤ 32	≥ 22

10-11 YEAR OLDS (n=133)

	TOTAL CORRECT	TOTAL COMMISSIONS
NORMAL	42 - 45	≤ 3
BORDERLINE	38 - 41	4 - 15
ABNORMAL	≤ 39	≥ 16

12-16 YEAR OLDS (n=154)

	TOTAL CORRECT	TOTAL COMMISSIONS
NORMAL	43 - 45	≤ 3
BORDERLINE	40 - 42	4 - 11
ABNORMAL	≤ 39	≥ 12

Table 3. Threshold Table for the Distractibility Task.

7-14 YEAR OLDS (N=58)	TOTAL CORRECT	TOTAL COMMISSIONS
NORMAL	28 - 45	≤ 6
BORDERLINE	6 - 27	7 - 35
ABNORMAL	≤ 5	≥ 36

Table 4. Time Analysis of Traditional Evaluations for ADD/Hyperactivity.

Number of cases evaluated	31
Mean number of hours spent:	
Administering Tests	3 - 5
Conducting Classroom Observations & Teacher Meetings	2 - 3
Interviewing Parents	1 - 2
Scoring Tests	1 - 2
Writing Reports	3 - 4
Total time expended (in hours)	10 - 16

Table 5. Accuracy of traditional evaluations using GDS scores as criteria.

	ADD	NON-ADD
HITS	13	6
MISSES	0 (FALSE NEGATIVES)	12 (FALSE POSITIVES)

Table 6. Percentage of children with ADD/Hyperactivity in self-contained classrooms

(N=29)

	n	ABNORMAL DELAY	ABNORMAL VIGILANCE	BOTH TASKS ABNORMAL
SLD	13	2 (15 %)	5 (38 %)	2 (15 %)
EMH	7	2 (29 %)	3 (43 %)	1 (14 %)
EH	9	5 (56 %)	6 (67 %)	4 (44 %)
TOTAL	29	9 (31 %)	14 (48 %)	7 (24 %)

(Psychopharmacology Bulletin)

Microprocessor-Based Assessment of Attention Deficit Disorders

Dissatisfaction with formulating a diagnosis of ADD/Hyperactivity based almost entirely upon clinical judgment or the perception of raters has spawned the development of behavior-based assessment procedures. These efforts typically involve administration of the Continuous Performance Test (Rosvold, Mirsky, Sarason, Brasome,Jr., & Beck, 1956), or related measures of attention and self-control (Davenport, 1972; Doyle, Anderson & Halcomb, 1976; Hiscock, Kinsbourne, Caplan & Swanson, 1979; Margolis, 1972). Although bulky and expensive electromechanical devices had previously been required, researchers are now programming microcomputers to administer these laboratory tasks to children (Conners, 1980; Klee & Garfinkel, 1983).

Use of the microcomputer has certain advantages, including flexibility of administration, the ability to store multiple data points, and, if a microcomputer is already available, cost-effectiveness. This approach, however, does have drawbacks for the researcher and, in particular, the clinician. The size of even portable microcomputers can make testing in multiple sites cumbersome. It is perhaps for this reason that standardization samples for software-driven programs tend to be, at best, limited in number and breadth. The transport of disk drives, monitors, keyboards, and the computer, itself, from location to location often discourages use of the procedure in other than a single research or clinic setting.

Another potential limitation of microcomputer-based testing stems from concerns surrounding the reliability of administration. Computer monitors vary in the size, shape, intensity and color of displayed characters. Each brand of computer presents to the child a different keyboard, and accepts different types of peripherals, such as "joysticks, which are often used as manipulanda. Without control over such critical factors as the amount of pressure necessary to actuate a switch, or the accuracy of timed sequences, repeatable administration across multiple microcomputers cannot be assured.

Finally, ADD/Hyperactive children, despite a reputation for academic underachievement, routinely display a fine facility for disassembling delicate equipment. Unless the examiner is immediately present, the highly impulsive child will inevitably stick a finger in disk drives, unplug cables, adjust monitors, or in some other fashion interrupt standard administration of the task. Efforts to secure the microcomputer are often unsatisfactory because they usually involve additional hardware, cabling, and expense.

An alternative approach to testing attention and self-control has been developed by the author (Gordon & McCLure, 1983; Gordon & McClure, 1984). The goal of this project has been to establish a practical, reliable, and well-standardized procedure available to both researchers and clinicians. The Gordon Diagnostic System (GDS) is a microprocessor-based, portable unit which, without an external microcomputer, allows for the administration of multiple tasks. The Delay Task, based upon a DRL operant schedule requires the child to inhibit responding in order to gain a reward. Specifically, the child is instructed to press a button, wait, and then press the button again. If s/he refrains from responding for at least six seconds, a light flashes and a reward counter increments. If the child responds before the interval elapses then the timer resets and no reward points are recorded. The GDS Delay Task generates three major scores: the number of responses (button presses), the number of correct responses (i.e. Correct) and the Efficiency Ratio which represents the percentage of correct responses.

The GDS also contains the Vigilance Task, which is a version of the Continuous Performance Test (Rosvold, et al., 1956). The child is presented with a series of digits on the front display and is told to press the button every time a "9" appears that had been preceded by a "1". The GDS records the number of correct presses, the number of times the child fails to press the button when the "1/9" combination appears (i.e. errors of omission), and the number of extraneous button presses (i.e. errors of commission). For the testing of younger children, the GDS contains a "1" mode, which requires the subject to press the button every time a "1" appears. The same performance measures are recorded. X

Although normative data were gathered using standard settings for task parameters, the design allows the user to select a wide range of parameters. This feature enabled the modification of parameters for the testing of adults as well as very young children.

Because the GDS is controlled by a programmable microprocessor, additional tasks can be explored. One such enhancement being studied is a version of the Vigilance Task which assesses the impact of distraction on a child's ability to sustain attention. This Distractibility Task is identical to the standard Vigilance Task except that random digits flash at random intervals on the outer positions of the LED display. The subject is still required to press the blue button when a "9" comes right after a "1". The only difference is that numbers flash on either side of the center (i.e. relevant) digit. Other enhancements under development are designed to evaluate attention across sensory modalities.

Even though the GDS administers tasks independently of a microcomputer, it can communicate with external hardware via an RS232C communications port. Software is available which allows for the direct transmission of GDS data to a microcomputer where it can be tabulated, graphed, stored, and compared to normative data. Connection of the GDS to a microcomputer enables collection of some ancillary data which cannot be extracted from the stand-alone unit.

The portability and ease-of-operation of the GDS has allowed for large-scale data collection. The standardization sample is comprised of 900 boys and girls from 3 to 16 years of age (Gordon & Mettelman, 1985). An additional 1100 hyperactive and nonhyperactive protocols from various subject populations, including the deaf, blind, Spanish-speaking, emotionally disturbed, and learning disabled, have also been gathered.

A series of validational studies has shown that these game-like tasks differentiated accurately between hyperactive and nonhyperactive children from both outpatient, and day treatment settings (Gordon, 1979; McClure & Gordon, 1984). In a sample of school-referred children, the GDS distinguished children with ADD from those classified as reading disabled, overanxious, and normal (Gordon & McCLure, 1983). Over 20 university and medical center research sites are currently conducting investigations involving the GDS. Most studies concern the effectiveness of the GDS for determining drug responsivity and for evaluating the success of pharmacotherapy (Atkinson, 1985; Barkley, 1985; Rapport, 1985). Others are investigating the use of the GDS for evaluating children with known brain damage, preschoolers considered "at-risk" for impulsive behavior, and the relationship between GDS scores and observational measures of classroom behavior. The GDS is also being used by school systems and pediatric practices across the country for the clinical evaluation of ADD/Hyperactivity.

COMMENTS AND CONCLUSIONS

A chronic impediment in the field of ADD/Hyperactivity has been the paucity of universal criteria and procedures for subject selection and treatment monitoring. Even in the realm of behavioral assessment, where truly creative approaches to testing children have been developed, procedures have generally not been sufficiently practical or well-standardized to allow for acceptance by the majority of professionals (Loney, 1981). As such, most continue to rely on rating scales or tests, such as the WISC-R (Wechsler, 1974), which have been shown to be of limited usefulness in differentiating between groups (Douglas, 1972). Furthermore, comparison of studies across research programs has often been hampered by idiosyncratic measures or sets of criteria.

The GDS represents an effort to establish a standard procedure for evaluating certain aspects of attention and self-control. While it is more costly and, in certain respects, less flexible than software-driven approaches, its portability, ruggedness, and extensive base of normative data offer the user greater assurance of reliable administration and valid interpretation. The growing acceptance of the GDS within the professional community also permits greater comparison of data across studies.

It must be emphasized, however, that the GDS was never intended as a divining rod for ADD/Hyperactivity. This is a complex disorder which represents an array of subgroups and interactive diagnostic dimensions. The GDS is seen as an important tool to be used only in conjuction with other selection criteria and clinical judgment.

References

Atkinson, A.W., Cohen, P.C. & Kelly, P.C. (1985, April). Attention deficit disorder: The effects of ritalin on self-esteem a comparison of ACTeRS teacher scale, Conners' parent scale, and Gordon Diagnostic System in diagnosis and management. Paper presented at the American Academy of Pediatrics Meeting, Atlanta, Georgia.

Barkley, R.A. (1985,August). Assessment of stimulant drug responding in ADD-H children. Paper presented at the 93rd Annual Convention of the American Psychological Association, Los Angeles, California.

Conners, C.K. (1980). Continuous performance test [Computer program]. Kensington, Maryland: Behavioral Medicine, Inc.

Davenport, W. (1972). Vigilance and arousal: Effects of different types of background stimulation. The Journal of Psychology, 82, 339-346.

Douglas, V.I. (1972). Stop, look and listen: The problem of sustained attention and impulse control in hyperactive and normal children. Canadian Journal of Behavioral Science, 4, 259-282.

Doyle, R.B., Anderson, R.P. & Halcomb, C.G. (1976). Attention deficits and the effects of visual distraction. Journal of Learning Disabilities, 9, 59-65.

Gordon, M. (1979). The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive children. Journal of Abnormal Psychology, 7, 317-326.

Gordon, M. & McClure, F.D. (1983). The objective assessment of attention deficit disorder. Paper presented at the 91st Annual Convention of the American Psychological Association, Anaheim, California.

Gordon, M. & McClure, F.D. (1984). Assessment of attention deficit disorders using the Gordon Diagnostic System. Paper presented at the 92nd Annual Convention of the American Psychological Association, Toronto, Canada.

Gordon, M. & Mettelman, B.B. (1985). Threshold tables for the Gordon Diagnostic System. (Available from Clinical Diagnostics, Inc., 300 E. Mineral Avenue, Suite 6, Littleton, Colorado 80122).

Hiscock, M., Kinsbourne, M., Caplan, B. & Swanson, J.M. (1979). Auditory attention in hyperactive children: Effect of stimulant medicaiton on dichotic listening performance. Journal of Abnormal Psychology, 88, 27-32.

Loney, J. (1981). Evaluating treatments for childhood hyperactivity: Some methodological considerations. In K.D. Gadow & J. Loney (Eds.) Psychosocial aspects of drug treatment for hyperactivity (pp. 77-103). Boulder, Colorado: Westview Press.

Klee, S.H. & Garfinkel, B.D. (1983). The computerized continuous performance task: A new measure of inattention. Journal of Abnormal Child Psychology, 11(4), 489-496.

Margolis, J.S. (1972). Academic correlates of sustained attention. Unpublished thesis, University of California, Los Angeles.

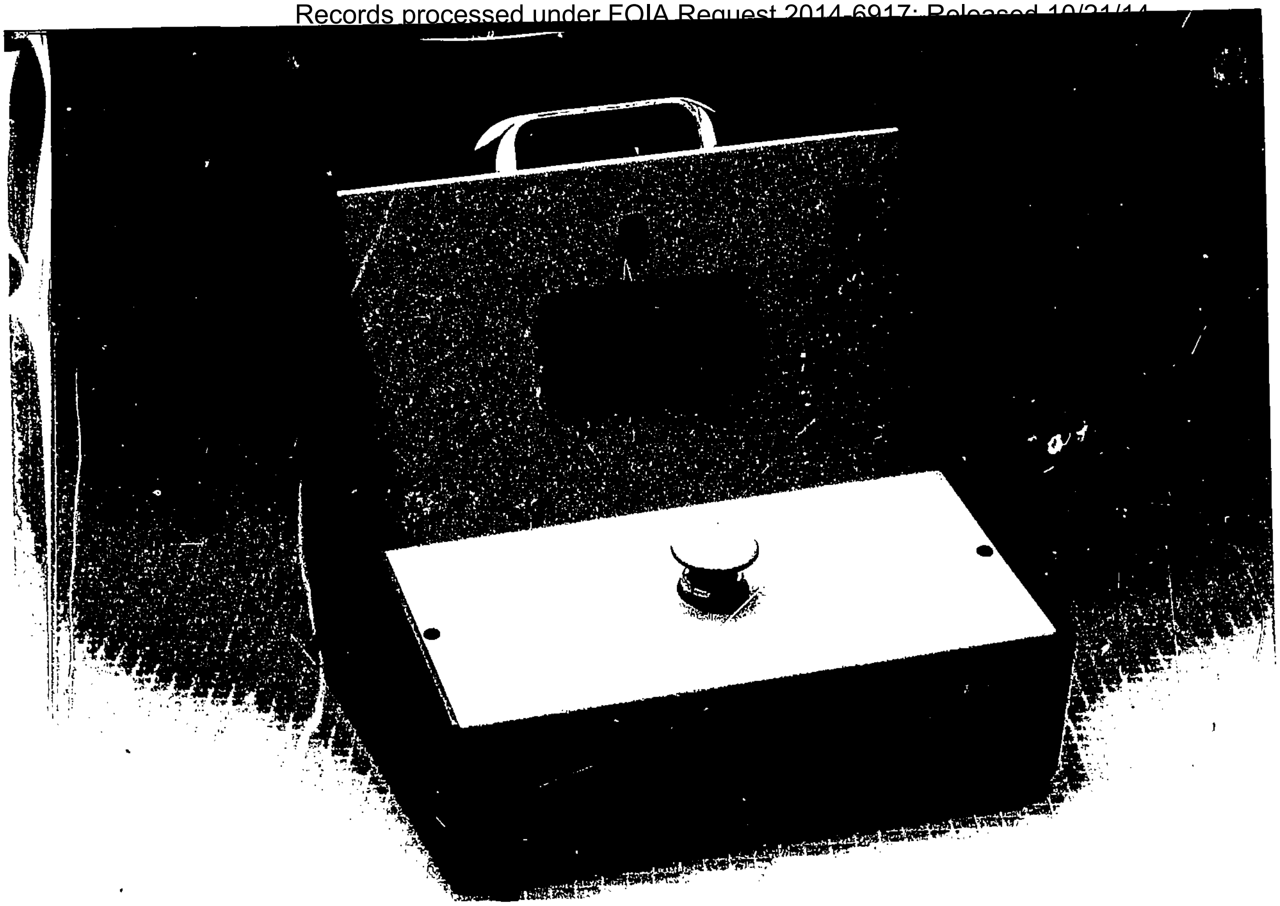
McClure, F.D. & Gordon, M. (1984). The performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. Journal of Abnormal Child Psychology, 12(4), 561-572.

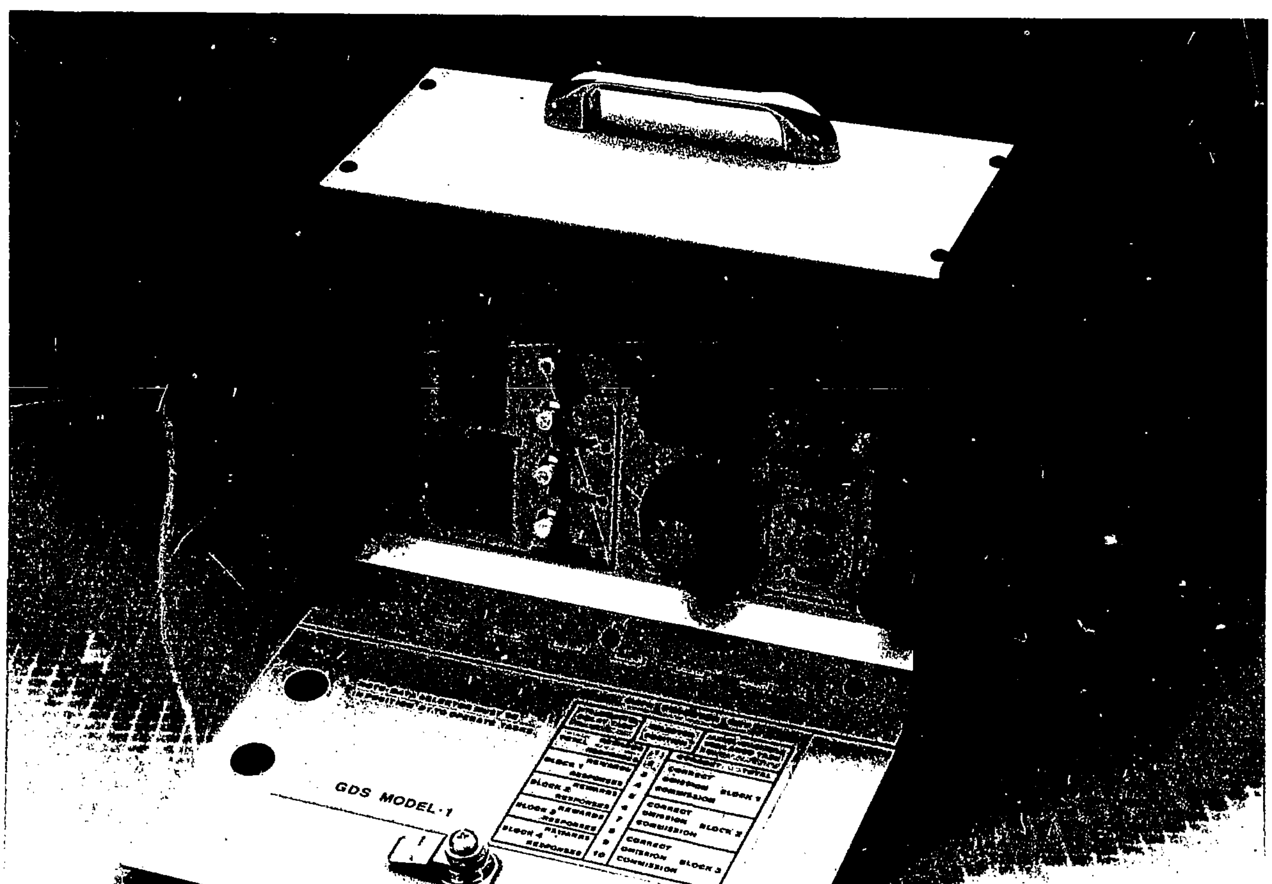
Rapport, M.D. (1985, August). Comparing classroom and clinic measures of ADD: Dose response effects. Paper presented at the 93rd Annual Convention of the American Psychological

Association, Los Angeles, California.

Rosvold, H.E., Mirsky, A.F., Sarason, I., Bransome, Jr., E.D. & Beck, L.H. (1956). A continuous performance of brain damage. Journal of Consulting Psychology, 20, 343-350.

Weschsler, D. (1974). Weschsler intelligence scale for children-revised. New York: Psychological Corporation.




GDS MODEL·1
CORRECT OMISSION COMMISSION BLOCK 2
CORRECT OMISSION COMMISSION BLOCK 3

K854903

- New Address -
300 E. Mineral Avenue. #6
Littleton, Colorado 80122
(303) 795-0438

Clinical Diagnostics, Inc.

November 25, 1985

Food and Drug Administration
Bureau of Medical Devices
Document Control Center (HFK-20)
8757 Georgia Avenue
Silver Spring, Maryland 20910

Re: 510K Notification

Attention: Document Control Clerk

This is to notify you of the intention of Clinical Diagnostics, Inc. to manufacture and market the following device:

Classification Name: Not Known

Common/Usual Name: Psychological Testing Device

Proprietary Name: Gordon Diagnostic System
GDS Data Analysis Program

Establishment Registration Number: 1721005

Classification: We are not aware that this device has been classified by FDA.

Performance Standard: None established under Section 514

Labeling/Promotional Material: Photographs and copies of promotional literature and manuals are enclosed.

Substantial Equivalence: We are not aware of any devices exactly like the Gordon Diagnostic System (GDS) and the GDS Data Analysis Program (GDSDAP). The two tasks performed in the GDS are based on psychological testing instruments developed for research. The Continuous Performance Test (CPT; Rusvold and Mirski, 1956) is the basis for the GDS Vigilance Task. The Differential Reinforcement of Low Rate Responding (DRL; Weisberg and Tragakis, 1967) is the basis for the GDS Delay Task.

The Gordon Diagnostic System Data Analysis Program is software to collate and analyze data produced by the GDS. It is not necessary for the operation of the GDS but serves as an accessory to facilitate data handling.

Respectfully submitted,

Bruce G. Meyer, Jr.
President

1536 COLE BOULEVARD, SUITE 350, GOLDEN, CO 80401
(303) 232-7129
(800) 521-4503

Clinical Diagnostics, Inc.

ABOUT THE GORDON DIAGNOSTIC SYSTEM

The Gordon Diagnostic System (GDS) is a tool to aid your diagnosis of children suspected of Attention Deficit Disorder (ADD). The instrument does this objectively by measuring two concerns of ADD: impulsivity and attentiveness. Game-like tasks, the Delay Task based on a DRL schedule, and the Vigilance Task based on a CPT schedule, result in recorded data that is compared to normed threshold tables.

DELAY TASK

The Delay Task is a self-paced task which measures the child's ability to surpress or delay impulsive behavioral responses. The task embodies a variety of processes, such as cognitive skills, behavioral demands, and motivational factors.

During the Delay Task, the child is instructed to press the button, wait a while, then press the button again. If the button is pressed too soon, no reward (point) will be obtained. Thus, the child must develop a method of determining the minimum time he must refrain from responding to obtain a point. He must develop a strategy for estimating each interresponse interval throughout the task.

The delay strategy employed involves a behavioral skill. Most nonimpulsive children use a "covert" strategy. That is, some form of internal, cognitively-oriented delaying procedure is employed. Usually, this is counting silently. Experience has shown that this method is quite effective, and children who use this method are quite adept at refraining from impulsive responses. Also, children who employ more "overt", or behaviorally manifested strategies, tend to perform much less efficiently on this task.

The Delay Task reflects the child's overall motivation to "do well". Since the task demands the application of cognitive and behavioral skills, the child must be reasonably motivated to maintain good performance throughout the task. The desire to gain control over a situation and the willingness to persevere with a rote task, are important factors in successful coping.

The microprocessor records time and responses. Standard parameters of a 6 second delay and four 120 second blocks are normed on approximately 1,000 subjects. Data recorded is transferred to a patient record form and compared to threshold tables. The Delay Task places maximum demands upon the ability to delay or refrain from emitting nonreinforced responses.

Clinical Diagnostics, Inc.

ATTENTION DEFICIT DISORDER/HYPERACTIVITY
BACKGROUND INFORMATION

PREVALENCE:

- 3-5% of school-aged population are affected. According to U.S. census figures, there are 30-40 million school-aged children.
- 30-40% of all children referred for professional help due to behavior problems come with a presenting complaint associated with ADD/Hyperactivity. By far, ADD is the most "popular" behavior disorder of childhood.
- 20% of adopted children are referred for symptoms of ADD/Hyperactivity.

DEMOGRAPHICS:

- Ratio of boys to girls averages 6:1.
- Highest incidence in lower socioeconomic status groups (SES) (highest in lower-middle class).
- IQ's are generally in the average range.

TYPICAL CHARACTERISTICS:

- Impulsivity.
- Inattentiveness.
- Restlessness.
- Academic Underachievement.
- Social Skill Deficits:
  - -Immaturity
  - -Egocentricity
  - -Poor regard for consequences of behavior
  - -Poor awareness of self and others
  - -Poor peer acceptance
- Emotional Immaturity.
- Exaggerated Emotional Reactions.
- Low Frustration Tolerance.
- Physical Findings:
  - -Motor Incoordination
  - -Enuresis (Bed Wetting) 30% of cases
  - -Encopresis (Soiling ones pants) 40% of cases
  - -Increased Incidence of Allergies
  - -Increased Incidence of Upper Respiratory Infection (URI)/cold and Otitis Media (middle ear infection)
  - -Increased Incidence of Minor Physical Anomalies
  - -Encephalogram (EEG) -- No Consistent Findings

OUTCOME:

- o Motor overactivity diminishes.
- o Chronic delay in problem areas.
- o Continued social difficulties.
- o Peer problems.
- o Increased association with children with conduct disorders.
- o Increased delinquency.
- o Poor school achievement.
- o Increased grade repetition.
- o Increased failure to finish school.
- o Decreased college attendance.

EMPLOYMENT DIFFICULTIES:

- o Not rated as abnormal by employer, but lower employer satisfaction.
- o Attains lower employment level.
- o Increased job and residence changes.
- o Emotional difficulties.
- o Increased depression.
- o Lower self esteem.
- o Lower frustration tolerance.
- o Increased hysteria in female hyperactives.

PREDICTORS OF OUTCOME:

- o Low IQ.
- o Low SES.
- o Presence of aggression/conduct disorder.
- o Degree of peer acceptance.
- o Degree of parent psychopathology.
- o Response to earlier treatment.

FAMILY CHARACTERISTICS:

- o 33-60% of parents with ADD child display similar symptoms.

Fathers:
- o Increased incidence of ADD/Hyperactivity symptoms.
- o Increased alcohol abuse.
- o Increased conduct problems.
- o Increased depression.
- o Increased desertion of family.
- o Increased difficulty maintaining consistent employment.

Mothers:
- o Increased depression.
- o Increased incidence of family discord/divorce (especially in lower SES).
- o First pregnancy at earlier age than common.
- o Increased alcohol abuse.
- o Increased stress in parenting situations.

Siblings:
- o Increased incidence of ADD/Hyperactivity symptoms.
- o Increased conduct problems.
- o Increased learning disabilities.

ASSESSMENT PROCEDURES:

Parents:
- o Careful Interview. (Review of symptoms; development history; social history; look for signs of parental depression and other effects of child's symptoms on parents).
- o Conners Parent Checklist or Achenbach Child Behavior Check List (CBCL).
- o Home Situations Questionnaire.
- o Locke-Wallace Marital Adjustment Survey.

Child:
- o Interview.
- o Gordon Diagnostic System (GDS).
- o Achievement Test.
- o Matching Familiar Figures Test (optional).
- o Review of school testing.

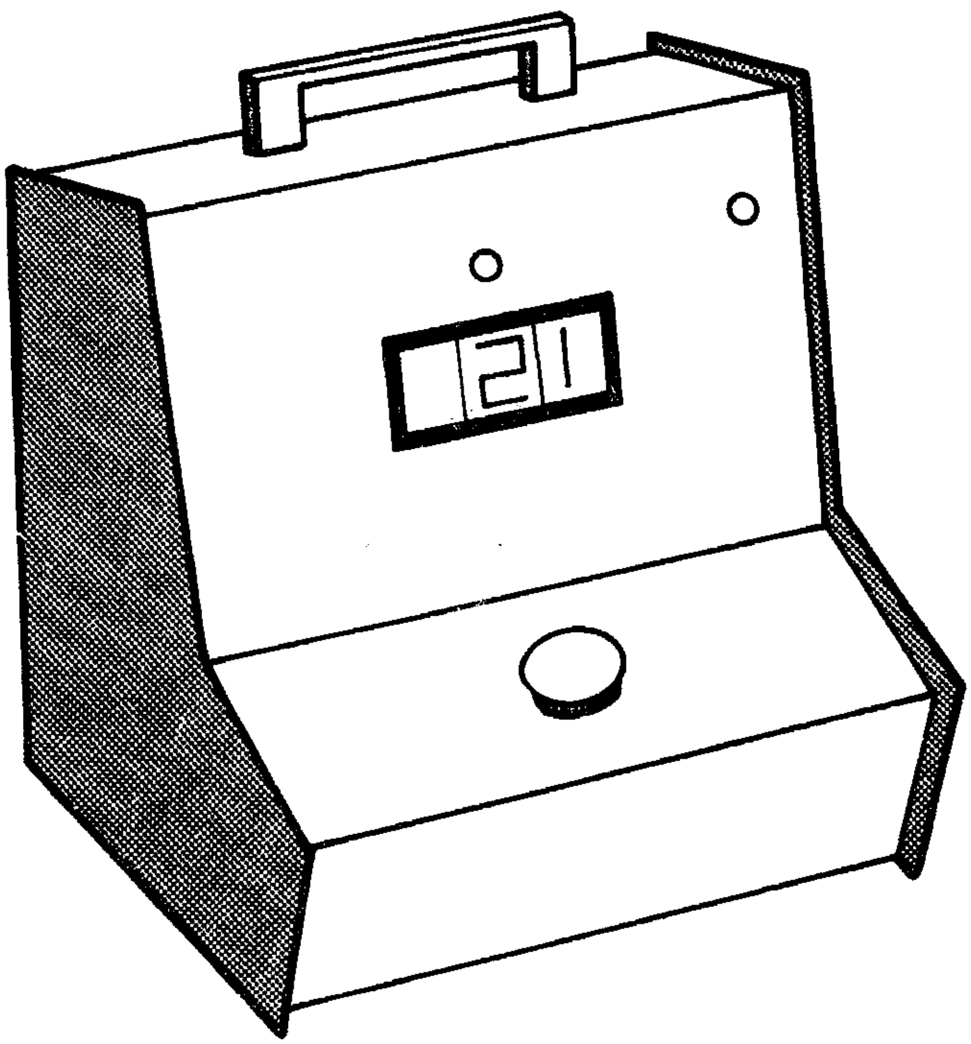
School:
- o Discussion with teachers.
- o Conners Teacher Rating Form or Achenbach Teacher Report Form.
- o Kendall-Wilcox Self-Control Questionnaire.

MEDICATION TREATMENT:

- o Ritalin most commonly prescribed.
- o Typical dosage is 0.3mg - 0.5mg/kg.
- o Should not be prescribed if presence of tics.
- o Improves conduct at home and school.
- o Does not improve academic achievement (some controversy here).

THE GORDON DIAGNOSTIC SYSTEM

INSTRUCTION MANUAL



CLINICAL DIAGNOSTICS INC
1536 Cole Blvd., Golden, Colorado 80401
Suite 350 • 303-232-7129

247

Introduction

These instructions provide all necessary information to operate the GDS successfully. Although the device has been designed to facilitate proper administration, **read and follow the instructions carefully!!**

The Gordon Diagnostic System (GDS) is a portable electronic device designed to assess deficits in attention and impulse control in children. It has been developed for use by clinicians as an aid in the diagnosis of attention deficit disorders with or without hyperactivity, as well as some forms of learning disabilities. Research has shown that the GDS yields important information about a child's functioning and can differentiate accurately between youngsters classified as hyperactive and nonhyperactive.

The GDS administers two tasks:
(1) The Delay Task measures a child's ability to refrain from pressing in order to win points and;
(2) The Vigilance Task assesses how well a child sustains attention over a long period of time.

Both qualities -- the ability to delay and the ability to maintain attention -- are considered by authorities to be the main areas where children diagnosed hyperactive run into trouble. Such children find it very difficult to stop, look, listen and think before they try to solve a problem. They are impulsive and inattentive, just the traits which the GDS measures. (For further information refer to the GDS Interpretive Guide).

Getting Started...

Open the rear panel with the key provided. (Controls for the GDS are kept locked to prevent children from tampering with them during the operation of the tasks.) Plug the GDS into a 120 volt, 60 Hz grounded socket and turn on the main power switch, located in the rear panel. The front display should light up and you should also hear a tone. If the display does not light, check to make sure that the plug is firmly in the socket.

NOTE: If you want to turn the GDS quickly off then on again, you must wait a few seconds before pushing the switch back to the **"ON"** position. If you do not wait long enough, the tone will not sound and you will have to turn the machine off, wait three seconds, and turn it on again.

248

Delay Task

The Delay Task, which has traditionally been administered before the Vigilance Task, requires the child to wait a set period of time (called the Delay Interval) before pressing the blue response button. If the subject has waited long enough, the Reward Light will shine and the Reward Counter on the front panel will increment. If the child presses the button before the Delay Interval has elapsed then no point is earned, the light stays off and the Delay timer resets.

The GDS records Delay Task performance over four successive time Blocks. You can select the length of the Blocks (1-999 seconds) and the Delay Intervals (1-99 seconds) independently for each Block. These Blocks are internal recording units to monitor the child's performance along the session. The child is not shown when one Time Block stops and another begins.

Standardization data for the Delay Task were gathered using a 6-second Delay Interval and identical time Blocks of 120 seconds each (Standard Parameters). To program parameters into the GDS microprocessors:

1. Set the Task Selector Switch to **DELAY TASK.** (The position of the Mode Selector is irrelevant to this task).

2. Set the **INTERVAL** thumbwheel to the number of seconds you want the child to wait between each press of the button. The Standard Parameter is 6 seconds.

3. Set the **BLOCK LENGTH** thumbwheel to the desired number of seconds. The Standard Parameters call for **BLOCK LENGTHS** of 120 seconds each.

4. To load the parameters into the GDS microprocessor, press the **BLOCK 1** button of the Block **PROGRAMMER.** This communicates to the GDS the parameters you desire for Block 1. If you want to enter these same parameters for the other three blocks, press the Block **PROGRAMMER** buttons for Blocks 2, 3, and 4. Therefore, if you want the identical **DELAY INTERVAL** and **BLOCK LENGTH** for all four blocks, set the thumbwheels to the desired number of seconds and press the four Block **PROGRAMMER** buttons. A tone should sound each time a **PROGRAMMER** button is depressed.

5. If you want to set different parameters for each block, change the thumbwheels to the desired settings and press the Block Programmer buttons to load the parameters for a particular block.

NOTE: When the power to the GDS is turned on, it automatically reads and loads the parameters showing on the thumbwheels into all Blocks. If those are the desired settings, you can start at step #6.

NOTE: If you make an error in entering parameters, simply reset the thumbwheels and press the appropriate Block Programmer button(s).

6. **SET THE DATA SELECTOR TO POSITION #1. THE TASK WILL NOT OPERATE UNLESS THE DIAL IS IN THE FIRST POSITION.**

7. The GDS is now ready to begin the Delay Task. Subjects have traditionally been given the following instructions:

   You're going to play a game in which you will get a chance to win a lot of points. Do you see this light (pointing to the small red light)? Every time you make this light go on you'll earn a point and this counter (pointing) will keep track of how many points you've won. At the end of the game we'll see how many points you've earned. Now, to make the light go on all you have to do is push this blue button (pointing), and wait a little while before pressing it again. You just press this blue button, wait a while, then press it again. If you press it again too soon, though, you will not get a point, the light won't go on, and you'll have to wait a while before you can press it to get another point. But if you push the button, wait a while, then push it again, you'll get a point every time.

Delay Task (Cont.)

8. At this time you will want to answer any questions and make sure that the child understands the instructions.

9. The examiner then says:

   I'll sit here and wait until you're done. You'll know when the game is over because the green light will come on again. Please let me know when the green light comes on. If you have any questions or want to talk about the game, I want you to wait until after the game is over, and we can discuss it then.

10. When ready to begin the task, press the **RESET** button and then the green **START** button. You can do this either before locking the back flap or after by pushing the buttons with a pencil through the holes. You should hear the click of a button as each is depressed. If you hear the tone the task has begun and you should tell the child that the game has started. If you do not hear the tone either the **DATA SELECTOR** dial was not set to position #1 or you did not press the buttons hard enough.

11. Now that the task has begun, you should sit behind and somewhat to the side of the child so as not to be a distraction. Conversation should be discouraged. As with any testing situation, observations of the child's approach to the task can be highly informative.

12. When the task is over a tone sounds and the green Game Over light will come on. You may find it helpful to ask the child questions about his or her approach to the task. Subjects have typically been asked, "What did you do to help you wait?" and "Did you find it easy or hard to play the game?".

13. At this point you can put the unit aside and go on to the other tests. The data will be stored **as long as the unit is turned on** and plugged in. If you wish to administer the Vigilance Task, you must first record the Delay Task data.

14. **TO RECORD DATA** open back flap with the key. By turning the **DATA SELECTOR** to each position you will see on the **DATA DISPLAY** the total number of rewards (times the button was pressed after waiting long enough) and responses (times the button was pressed) for the whole session, as well as for each of the 4 Time Blocks. The key for the **DATA SELECTOR** positions is printed on the inside of the back flap and are as follows:

Data Selector Position	Data Displayed
1	Total Rewards
2	Total Responses
3	Block 1 Rewards
4	Block 1 Responses
5	Block 2 Rewards
6	Block 2 Responses
7	Block 3 Rewards
8	Block 3 Responses
9	Block 4 Rewards
10	Block 4 Responses

Make sure the Data Selector is set to position #1 when you are finished recording the data.

15. Once you have finished writing down the results you can either turn the unit off, move on to the Vigilance Task, or readminister the Delay Task. If you turn the power off all data is erased from the GDS microprocessor.

NOTE: If you wish to clear the front display, press the **RESET** button. If you have not turned off the power and want to readminister the Delay Task with the same parameters, simply press **RESET** and **START** to begin the task.

250

Delay Task

The Delay Task, which has traditionally been administered before the Vigilance Task, requires the child to wait a set period of time (called the Delay Interval) before pressing the blue response button. If the subject has waited long enough, the Reward Light will shine and the Reward Counter on the front panel will increment. If the child presses the button before the Delay Interval has elapsed then no point is earned, the light stays off and the Delay timer resets.

The GDS records Delay Task performance over four successive time Blocks. You can select the length of the Blocks (1-999 seconds) and the Delay Intervals (1-99 seconds) independently for each Block. These Blocks are internal recording units to monitor the child's performance along the session. The child is not shown when one Time Block stops and another begins.

Standardization data for the Delay Task were gathered using a 6-second Delay Interval and identical time Blocks of 120 seconds each (Standard Parameters). To program parameters into the GDS microprocessors:

1. Set the Task Selector Switch to **DELAY TASK.** (The position of the Mode Selector is irrelevant to this task).
2. Set the **INTERVAL** thumbwheel to the number of seconds you want the child to wait between each press of the button. The Standard Parameter is 6 seconds.
3. Set the **BLOCK LENGTH** thumbwheel to the desired number of seconds. The Standard Parameters call for **BLOCK LENGTHS** of 120 seconds each.
4. To load the parameters into the GDS microprocessor, press the **BLOCK 1** button of the Block **PROGRAMMER.** This communicates to the GDS the parameters you desire for Block 1. If you want to enter these same parameters for the other three blocks, press the Block **PROGRAMMER** buttons for Blocks 2, 3, and 4. Therefore, if you want the identical **DELAY INTERVAL** and **BLOCK LENGTH** for all four blocks, set the thumbwheels to the desired number of seconds and press the four Block **PROGRAMMER** buttons. A tone should sound each time a **PROGRAMMER** button is depressed.
5. If you want to set different parameters for each block, change the thumbwheels to the desired settings and press the Block Programmer buttons to load the parameters for a particular block.

NOTE: When the power to the GDS is turned on, it automatically reads and loads the parameters showing on the thumbwheels into all Blocks. If those are the desired settings, you can start at step #6.

NOTE: If you make an error in entering parameters, simply reset the thumbwheels and press the appropriate Block Programmer button(s).

6. **SET THE DATA SELECTOR TO POSITION #1. THE TASK WILL NOT OPERATE UNLESS THE DIAL IS IN THE FIRST POSITION.**
7. The GDS is now ready to begin the Delay Task. Subjects have traditionally been given the following instructions:
   You're going to play a game in which you will get a chance to win a lot of points. Do you see this light (pointing to the small red light)? Every time you make this light go on you'll earn a point and this counter (pointing) will keep track of how many points you've won. At the end of the game we'll see how many points you've earned. Now, to make the light go on all you have to do is push this blue button (pointing), and wait a little while before pressing it again. You just press this blue button, wait a while, then press it again. If you press it again too soon, though, you will not get a point, the light won't go on, and you'll have to wait a while before you can press it to get another point. But if you push the button, wait a while, then push it again, you'll get a point every time.

251

Vigilance Task

The Vigilance Task has two modes. In the "1/9" mode a series of numbers appears on the front display and the child is required to press the button every time a "9" appears immediately after a "1". The GDS records the number of Correct presses, the number of times the child did not press the button upon the appearance of a "1/9" combination (Omissions), and the number of improper presses (Commissions). The "1" task is similar, but for children under six years of age or for those who cannot follow the "1/9" instructions. The child is told to press the button every time a "1" appears.

The GDS records performance on the Vigilance Task over **three** Time Blocks (unlike four for the Delay Task). Normative data was derived using three identical Time Blocks of 180 seconds each.

The Presentation Interval is the time from when a number first appears to the initial presentation of the next digit. A Presentation Interval of one second was the Standard Parameter for the standardization studies. Numbers always appear for two tenths of a second (i.e. 200 milliseconds) regardless of the Presentation Interval setting. Consequently, when the Interval is set at 1 second, the digit appears for .2 seconds and the blanking time (when nothing appears on the screen) is .8 seconds. If you chose an **INTERVAL** of 2 seconds, the number appears for .2 seconds with a blanking time of 1.8 seconds.

Procedures for setting parameters and recording data for the Vigilance Task are very similar to those for the Delay Task:

1. Set the Task Selector Switch to **VIGILANCE TASK** and the Mode Selector either to "1/9" or "1". Clear the front display by pressing the **RESET** button.
2. Set the **INTERVAL** thumbwheel to the desired number of seconds you want between presentation of the digits. An **INTERVAL** of 1 second was used to establish normative data.
3. Set the **BLOCK LENGTH** thumbwheel to the desired number of seconds. The Standard Parameters call for Block Lengths of 180 seconds for each of the three time Blocks.
4. To enter the parameters into the GDS for Block 1, press the Block 1 button of the **BLOCK PROGRAMMER.** If you want to enter these same parameters for the other two Blocks, press the **BLOCK PROGRAMMER** buttons for Blocks 2 and 3 (there is no Block 4 in the Vigilance Task). Again, the procedure for establishing identical Presentation Interval and Block Length parameters for all three blocks is to turn the thumbwheels to the desired settings and press the first three **BLOCK PROGRAMMER** buttons.
5. If you want to set different parameters for each block, change the thumbwheels to the desired settings and press the Block Programmer buttons to register the parameters for a particular block.
6. **THE DATA SELECTOR MUST BE SET TO POSITION #1. THE TASK CANNOT RUN UNLESS THE DIAL IS IN THE FIRST POSITION.**
7. Close and lock the back flap. The GDS is now ready to administer the Vigilance Task. Children have typically been given the following instructions:

   Now we are going to play a different game. This time you are going to see numbers flash on the screen and I want you to press the blue button only when you see a "1" and then a "9". If a "9" shows by itself without a "1" coming right before it then do not press the button. Only press the button if you see a "9" coming right after a "1".

   Children performing on the "1" Mode are told simply to press the button every time the number "1" appears. With very young children it is important to first check that they can recognize the number "1".
8. Before proceeding you will want to answer any questions and make sure that the child understands the instructions.

252

Vigilance Task (Cont.)

9. The examiner then says:

   I'll sit back here and wait until you've finished. You will know when the game is over when the green light comes on. If you have any questions or want to talk about the game, I want you to wait until the game is over and we can discuss it then.

10. To begin the game press the **RESET** and **START** buttons making sure you hear each click as it is depressed. The task will begin as soon as the **START** button is pushed and the tone sounds. You will want to tell the child to begin immediately. If you do not hear the tone then check to make sure the Data Selector dial is in position # 1 and then firmly press the **START** button.

11. While the child is performing on this task you should again sit behind and to the side.

12. The task has been completed when the tone sounds and the green Game Over light comes on. The data for the Vigilance Task will be be stored as long as the power is turned on and the **RESET** button has not been pressed.

13. At this point you can put the unit aside and go on to the other tests. The data will be stored **as long as the unit is turned on** and plugged in. If you wish to administer the Delay Task, you must first record the Vigilance Task data.

14. **TO RECORD DATA** open back flap with the key. By turning the **DATA SELECTOR** to each position you will see on the **DATA DISPLAY** the total number of Correct responses for the whole session as well as the Correct responses, Omissions and Commissions for the 3 time Blocks. The key for the **DATA SELECTOR** positions is printed on the inside of the back flap and is as follows:

Data Selector Position	Data Displayed
1	Total Correct
2	Block 1 Correct
3	Block 1 Omissions
4	Block 1 Commissions
5	Block 2 Correct
6	Block 2 Omissions
7	Block 2 Commissions
8	Block 3 Correct
9	Block 3 Omissions
10	Block 3 Commissions

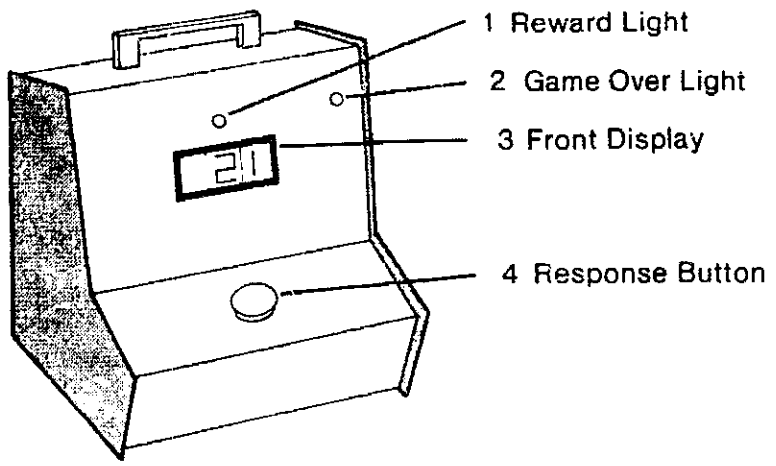
Make sure the **DATA SELECTOR** is set to position #1 when you are finished recording the data.

15. Once you have finished writing down the results you can either turn the unit off or repeat one of the tasks. If you turn the power off all data will be lost!

NOTE: If you wish to clear the front display, press the **RESET** button.

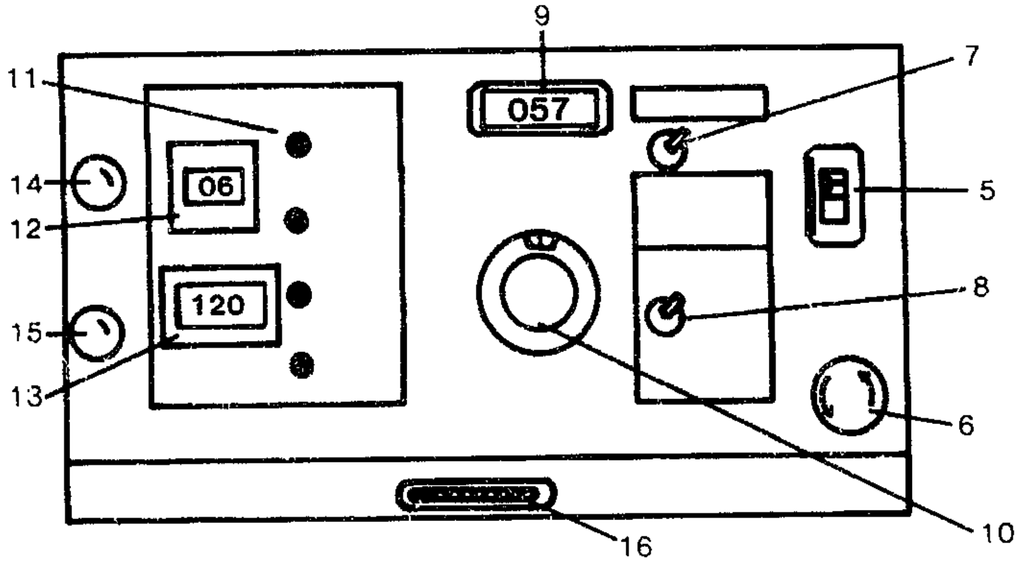
253

Description Of The GDS



Front View

1. Reward Light - Indicates rewarded response during Delay Task.
2. Game Over Light - Indicates end of Task.
3. Front Display - Counts rewards during Delay Task. Displays digit series during Vigilance Task.
4. Response Button - Pressed by child to perform Task.



Rear View

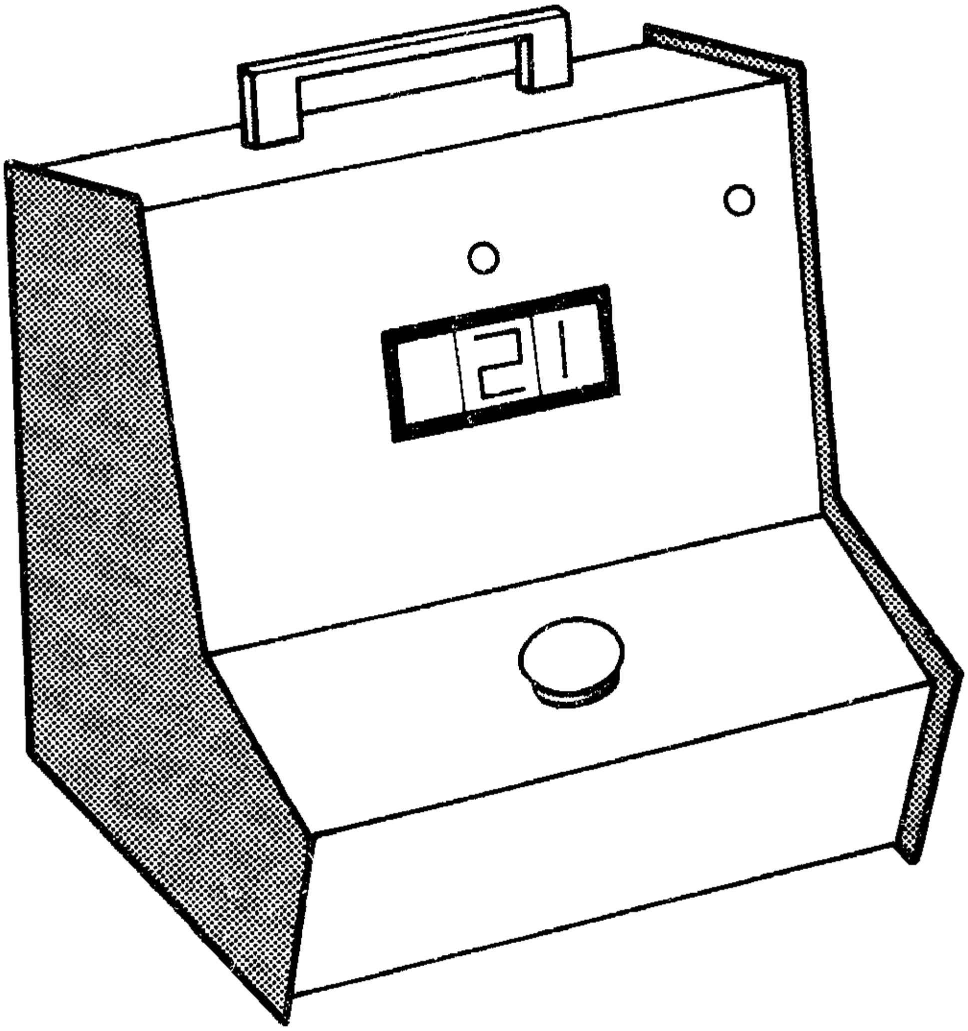
5. Power Switch - Turns power on and off.
6. Fuse - Protects GDS from damage due to mechanical difficulties.
7. Task Selector Switch - Allows selection of Delay or Vigilance Task.
8. Mode Selector - Allows for selection between "1" and "1/9" version of Vigilance Task.
9. Data Display - Displays data from GDS performance.
10. Data Selector - Allows user to determine which data to display.
11. Block Programmer - Loads thumbwheel settings into GDS microprocessor.
12. Interval Thumbwheel - Sets Delay Interval for Delay Task and Presentation Interval for Vigilance Task.
13. Block Length Thumbwheel - Sets length of individual Time Block.
14. Reset Switch - Resets counters and clears displays.
15. Start Switch - Press to begin task.
16. RS-232 Communications Port - Enables GDS to link up with external computer terminals.

254

CLINICAL DIAGNOSTICS INC.
1536 Cole Blvd., Golden, Colorado 80401
Suite 350 • 303-232-7129

INTERPRETIVE GUIDE

THE GORDON DIAGNOSTIC SYSTEM



CLINICAL DIAGNOSTICS INC.
1536 Cole Blvd., Golden, Colorado 80401
Suite 350 • 303-232-7129

INTERPRETIVE GUIDE TO THE GORDON DIAGNOSTIC SYSTEM

Michael Gordon, Ph D.
S.U.N.Y. Upstate Medical Center, Syracuse, New York

F. Daniel McClure, Ph.D.
D.C. Wilson Neuropsychiatric Hospital, Charlottesville, Virginia

Ernest M. Post, M.D.
S.U.N.Y. Upstate Medical Center, Syracuse, New York

Acknowledgments

The authors wish to express their gratitude to Ms. Jennifer DeLisi who, with great competence and cheer, coordinated our current research programs and assisted in the preparation of this Interpretive Guide. We also appreciate the fine efforts of Mrs. Mary Brown, Mrs. Noreen Lannon, and Miss Jennifer Kagan, and the crucial support of Mr. Kevin Bacher of the Liverpool School District, Liverpool, New York. These acknowledgments would not be complete without a note of gratitude to our families who have endured much throughout the development of the GDS. We are particularly thankful for the expertise and tolerance of Dr. Wendy Evers Gordon.

Aug. 17, 1983

Table Of Contents

259

INTRODUCTION TO THE INTERPRETIVE GUIDE

What is commonly referred to as hyperactivity probably raises more concerns than any other childhood behavior disorder of our time. It is also one of the most confusing and controversial issues ever to emerge in the arena of child mental health. The discussion of hyperactivity has highlighted the philosophical, theoretical, and professional biases among educators, pediatricians, psychiatrists, and psychologists.

How Common Is Hyperactivity?

Conservative estimates place the incidence of hyperactivity at between five and ten percent of the entire national elementary school population (Miller, Palkes, & Stewart, 1973). It is frequently at the top of the list when parents and educators are questioned about child behavior problems; it is by far the most frequent reason for consultations with child psychiatrists, pediatricians, and child psychologists; and it is a problem of major proportions in our nation's educational system. In a study of the Des Moines, Iowa public schools (Stone, Spence, Wilson, & Gibson, 1969) it was found that 53% of the boys and 30% of the girls were identified by their teachers as having symptoms of hyperactivity. Further, a national news report (Maynard, 1970) revealed that between five and ten percent of the 62,000 elementary school students in the Omaha, Nebraska public school system were being given behavior modifying drugs after being identified as hyperactive. These findings suggest a problem of epidemic proportions, but also reflect the confusion and misunderstanding surrounding the concept of hyperactivity. They dramatically speak to the paucity of valid, reliable diagnostic methods currently available in this area. In fact, accurate estimates of the true incidence of the disorder are not yet possible due to confusion over terminology and glaring inadequacies in diagnostic methods.

How Has Hyperactivity Traditionally Been Defined?

Much of this confusion arises from use of the term "hyperactivity". It has been very loosely applied to a wide variety of children who share a few common features; there has been little consensus among professionals as to exactly what the term implies. Some interpret it solely as a descriptive term referring to a child's overall level of motor output; others view it as reflecting the presence of neurological damage or "minimal brain dysfunction"; some link the term to certain autonomic dysfunctions such as arousal level or autonomic responsivity; still others consider it a disorder of higher cognitive structures, such as those involved in "metacognitive strategies" for problem solving. Since hyperactivity is not a specific disease, like diabetes, the term has been used in a variety of individually determined ways. Usually, a professional's view of hyperactivity is primarily determined by his or her own personal value system, theoretical orientation, professional training, and general philosophical view of development. It is no surprise, then, that highly subjective criteria have routinely been employed in "diagnosing" this disorder. As a result, the term has quickly lost any specificity which it might have otherwise embodied, and has essentially evolved into a rather meaningless "catch-all" category for many problem children. This situation has prompted some people to feel that hyperactivity simply represents "those aspects of a person's behavior which annoy the observer" (Buddenhagen & Sickler, 1969).

In response to this ambiguity, the DSM III psychiatric nomenclature has adopted the term "Attention Deficit Disorder With or Without Hyperactivity (ADD)". This change reflects the growing emphasis on evaluating the quality rather than the sheer quantity of the child's behavior. (The ADD classification will be used whenever possible in this manual. However, for the sake of continuity with the existing professional literature, the term hyperactivity is occasionally employed).

How Has The Disorder Been Diagnosed?

Richard Gardner (1979) noted that many professionals have routinely diagnosed the disorder solely on the basis of parent and/or teacher impressions, without actually having observed firm diagnostic signs themselves. This strongly contrasts with the usual practice of professionals, especially physicians, who rarely accept any other diagnosis made by a lay person. This sad state of affairs takes on added significance when one considers that psychostimulant medication has long been the treatment of choice for these youngsters and is still often used when the "diagnosis" is made. Physicians have frequently been the subject of heated indictments in professional journals, as well as in the lay press, for loose and indiscriminate prescription of these medications for supposedly affected children. Unfortunately, there have not been any effective alternatives to the subjective, highly biased, and often grossly inadequate methods of diagnosis. These inadequacies have often lead to ill-conceived and frequently inappropriate forms of treatment.

Under these circumstances, the need for a valid and objective diagnostic methodology cannot be overstated. Accurate diagnosis and assessment are as essential to effectively treating those who need it as they are to avoiding unnecessary treatment. If the basic assessment is erroneous or inadequate, then the effectiveness of treatment is seriously compromised.

Assessment Of ADD/Hyperactivity: A Conceptual Approach

Our approach to the evaluation of behaviors classified as hyperactive or ADD rests upon the functional assessment of the critical dimensions of impulse inhibition and attention. Our experience and a large body of research literature dictate that the essential features of the disorder lie in the inability to suppress behavioral responding and in the inability to focus and sustain attention. These processes are obviously very much interrelated. For example, it is difficult to expect a child to inhibit impulses to "act" when he or she cannot maintain attention for any length of time. Deficits in the ability to focus and sustain attention are therefore one of many causes for responses which are impulsive and poorly organized. By the same token, it is difficult to expect a child who is highly impulsive for other reasons and who cannot delay his or her responding to be able to attend to a task for any length of time. Their behavioral impulsivity seriously interferes with their ability to attend. This constellation of attentional deficits and lack of impulse inhibition have been referred to as the "stop, look, and listen" dimensions (Douglas, 1972), and they are the focus of our approach to the disorder. To reiterate, the two fundamental features involved in ADD are:

(a) the inability to refrain from emitting impulsive, and poorly directed behaviors, and
(b) the inability to focus attention and/or, sustain it over time.

Deficits in these two areas produce a deeply-ingrained pattern of behavior which we call the "impulsive style". This impulsive style is at the heart of what brings the child into conflict with his environment. We consider the term "hyperactivity" actually to refer to the symptoms of an impulsive style. The behavior of impulsive children may vary greatly from child to child, depending on a myriad of internal and external factors. Some children's impulsivity may manifest itself primarily at the level of attention and concentration and not in actual physical behavior. This child would most likely be diagnosed as having "ADD Without Hyperactivity" and require intervention aimed at developing stategies for sustaining attention and vigilance. The impulsive style of other children may also express itself through physical routes, and these children are often diagnosed as having "ADD with Hyperactivity". They require additional strategies for adopting less physical means of mediation and often are prescribed stimulant medication.

Emphasis on the impulsive style leads us away from focusing on specific behaviors or clusters of behaviors in evaluating these children. Instead, we have come to look for accurate assessment of the underlying processes of attention and delay which are at the core of the disorder. We assume that children with these deficits, i.e., who have an impulsive style, are at a significant disadvantage in adequately coping with problems that confront them.

It should be clear that our approach to the assessment of this disorder is not bound to any particular theoretical framework or conviction about etiology. Instead, we seek to describe, in an objective and realistic manner, the child's functional abilities and disabilities with the expectation that careful evaluation will lead to a better understanding of these problems and more efficient treatment planning. Indeed, the Gordon Diagnostic System (GDS) was developed with these goals in mind. We believe that it represents an accurate, stable, and objective approach to the assessment of the impulsive style and related disturbances of attention.

This interpretive guide provides further information on the background of the GDS and offers suggestions for interpreting the results of testings obtained with the device. Data from our research studies and clinical experience are presented, as are standardization tables for your use. To avoid gender prejudice, we have alternated male and female references throughout the manual.

261

THE ASSESSMENT OF ATTENTION DEFICIT DISORDERS

Background

Clinicians, educators, and physicians have been in chronic need of reliable and valid techniques for the assessment of attention deficit disorders (ADD) and impulsivity in children. The instruments that are currently available for the diagnosis of ADD fall into three basic categories: those actually intended for the measurement of other aspects of behavior which have been adapted to assess impulsivity, those which are largely limited to laboratory investigations, and questionnaires.

Most prominent among the instruments originally designed for the clinical assessment of other behavioral domains is the Wechsler Intelligence Scale for Children-Revised (WISC-R). While selected subtests or groups of subtests from the WISC-R are widely used in the identification of hyperactivity, studies have shown that they do not discriminate accurately between groups of hyperactive and nonhyperactive children (Douglas, 1972). Moreover, the WISC-R assesses a wide range of abilities and traits that are largely unrelated to impulsivity (Douglas, 1972). Thus, while all hyperactive children may perform poorly on these measures, there are many nonhyperactive children (e.g. learning disabled, or those with visual-motor deficits) who may also have difficulties with them for reasons other than impulsivity.

Among those instruments best suited for laboratory investigations, the most effective in differentiating hyperactive from nonhyperactive children are the Continuous Performance Test (Rosvold, Mirsky, Sarason, Bransone, & Beck, 1956), Delayed Reaction Time Test (Cohen, Douglas & Morganstern, 1972), and the Matching Familiar Figures Test (Kagan, 1966). While the Continuous Performance Test and the Delayed Reaction Time Test have been demonstrated to be excellent diagnostic measures, their technological sophistication and prohibitive expense make them unsuitable for widespread clinical use. The MFFT is quick, easy to administer, and has been demonstrated to be diagnostically sound. However, its reliability and validity have been the subject of considerable controversy which persists despite revisions of the measure (Egeland & Weinberg, 1976).

There are a variety of other measures available such as the Gardner Steadiness Tester (Gardner, 1979), Human Figure Drawings (Machover, 1949), certain configurations of Rorschach indicies, and paper-and-pencil "vigilance" tasks. As a group, these tend to suffer drawbacks stemming from impracticality of administration, limited information yield, and often poor standards of validity and reliability.

Behavioral inventories or questionnaires have also enjoyed widespread use, and many, such as the Teacher Rating Scale (Conners, 1969), the Child Behavior Check List (Achenbach, 1978), and the Self-Control Rating Scale (Kendall & Wilcox, 1979), have demonstrated good discriminative validity. However, there are serious inherent drawbacks to such a format, including rater bias, the possibility of "halo" effects, item overlap, vague rating criteria, issues of basic validity, and the fact that questionnaires yield no direct behavioral data.

There is thus an obvious and compelling need for a more straightforward, clinically applicable, and precise measure of hyperactive child's impulsivity. The Gordon Diagnostic System (GDS) shows considerable promise as a device which could help fill this gap in delivery of clinical service. The GDS is an easily-administered, game-like task which seems to precisely measure levels of impulsivity without contamination by other factors such as intelligence or visual-motor skills. It provides objective data on a particular child's ability to inhibit behavioral responding. Through observation of the child's performance and from responses to several post-test inquiries, important insights can be gained into a child's problem-solving style. This information can have strong bearing on treatment recommendations and educational planning. Thus, the GDS is unique because it yields not only quantitative data, but also a wealth of pertinent behavioral information which is crucial to the diagnostic process.

History Of The GDS

The Gordon Diagnostic System is an outgrowth of research by the senior author, Dr. Michael Gordon. While serving as a research assistant in the physiological psychology laboratory at The Ohio State University, Dr. Gordon was involved in a study of rats with lesions to the septal area of the brain (Gittis & Gordon, 1977). Following the ablations, these rats and a matched control of non-operated rats were tested on a Differential Reinforcement of Low Rates (DRL) schedule. The septal animals did poorly on this schedule because they lost the ability to adequately inhibit responding. At this same time, the literature on hyperactivity was pointing to impulsivity and attentional disorders as the core symptoms of hyperactivity (Douglas, 1972). It was also claimed that hyperactive children suffered from damage to the septal area. It therefore seemed sensible to investigate the performance of children classified as hyperactive on a customized version of the DRL paradigm (Gordon, 1979). Because of the promising results, the technique was further refined so that the procedure could become available to practicing clinicians. Most importantly, considerable reduction in the size of the equipment was made possible through the use of integrated circuits.

Description Of The GDS

The GDS is a portable, electronic unit which contains two tasks. The Delay Task is based upon a Differential Reinforcement of Low Rates (DRL) behavior schedule and requires the child to inhibit responding in order to gain a reward. Specifically, the child is instructed to press a button, wait, and then press the button again. If he refrains from responding for at least six seconds, a light flashes and a counter increments. If the child responds before the interval elapses then the timer resets and no reward points are

262

recorded. The GDS Delay Task generates three major scores, the number of responses, the number of correct responses (i.e. Rewards) and the Efficiency Ratio which represents the percentage of correct responses.

The GDS also contains the Vigilance Task which is a version of the Continous Performance Test (Rosvold, et al., 1956). The child is presented with a series of digits on the front display and is told to press the button every time a "9" appears that had been preceded by a "1". The GDS records the number of correct presses, the number of times the child fails to press the button when the "1/9" combination appears (i.e. errors of omission), and the number of extraneous button presses (i.e. errors of commission). For the testing of children under 6 years of age, the GDS contains a "1" mode which requires the subject to push the response button every time a "1" appears. The same performance measures are recorded.

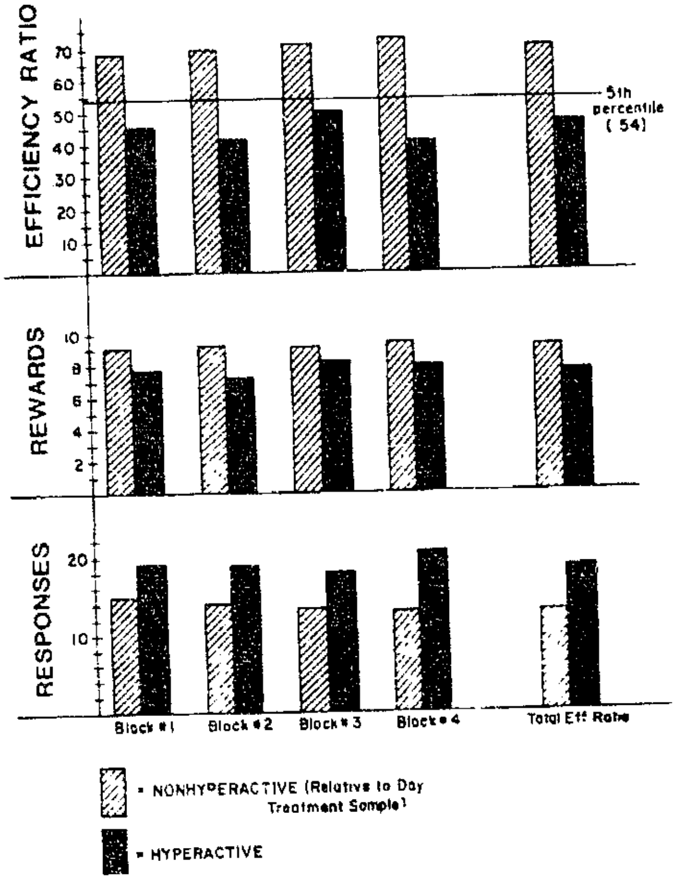
We have recently developed a GDS Data Analysis Program which enables a microcomputer to receive data directly from the GDS. The information can then be tabulated, graphed, and stored. The Data Analysis Program also allows measures of performance characteristics not available on the device itself (see "Ongoing Research").

Research Studies With The GDS

Data concerning the GDS Delay Task come from six studies conducted over as many years. The original experiment (Gordon, 1979) compared performances of 20 hyperactive and 20 nonhyperactive children selected from the patient roster of a child guidance clinic. There were highly significant differences between the two groups for all three GDS variables, i.e. Rewards, Responses and the Efficiency Ratio (Rewards/Responses). The children classified as hyperactive had profound difficulties in dealing successfully with the delay inherent in the task.

Since this initial work we have conducted several previously unreported studies on GDS performance. The original finding was recently replicated (McClure & Gordon, in press) with patients who had been placed in a day treatment program for seriously disturbed youngsters. Differences in GDS Delay Task scores between the 20 children classified as hyperactive and the 20 nonhyperactives in the comparison group were strikingly significant (p < .0001, see Figure).

DELAY TASK PERFORMANCE OF HYPERACTIVE VS. NONHYPERACTIVE SUBJECTS



To further judge the discriminative ability of the GDS, analyses were performed with groups formed on the basis of being identified as hyperactive or nonhyperactive by four criterion measures. These included the Teacher Rating Scale (Connors, 1969), the Self-Control Rating Scale (Kendall & Wilcox, 1979), the Teacher Report Form (Achenbach, 1980), and the Matching Familiar Figures Test - 20 (Cairns & Cammock, 1978). For subjects who were clearly identified as hyperactive (n = 8) or nonhyperactive (n = 8) on the basis of all four criteria, there was agreement between the GDS and the criterion measures in 15 cases (94%, Fisher's Exact Test, p = .01). When children were classified as hyperactive or nonhyperactive by **at least** three of the four criterion measures, there was agreement in 29 of 32 cases (91%, Fisher's Exact Test, p = .001).

Two additional studies appear to establish the validity of the GDS as a measure of impulsive/hyperactive behavior. In a group

263

of 31 nonhyperactive subjects. GDS performance was significantly correlated with 12 teacher ratings. It was also related to impulsivity as measured by the Bender-Gestalt Test and the Rorschach Inkblots. Analysis of data from the files of 32 outpatient children also indicate strong relationships between GDS performance and parent ratings of impulsive behavior.

To establish norms for the GDS Delay Task scores, a standardization study was recently completed in the public schools. The sample included a total of 192 boys and girls in three age groups: 6-7 yrs.; 8-9 yrs., 10-11 yrs.). It was determined that the mean Efficiency Ratio for the Delay Task (i.e. Total Rewards/Total Response was .84 ± .13 and that the scores were unaffected by age, sex, IQ and socioeconomic status. As with previous studies, performance remained consistent throughout the 8-minute session. In order to establish test-retest reliability, half the group was retested after 30-45 days, and it was found that there was a small learning effect that varied little among children.

Also completed recently was a study of 60 children who were classified by the school district as falling into one of four groups: attention deficit disorder, reading disabled, overanxious and normal. The GDS Delay and Vigilance Task successfully differentiated between the hyperactive group and the other three categories. It was also found that the correlations between the Delay Task Efficiency Ratio and the Vigilance Task scores fell in the .5 to .6 range of Pearson's r.

The Vigilance Task was first included in the GDS half-way through the standardization study. As such, there are now available norms for Vigilance Task performance. It was also used in the study of classified children and that data is currently being analyzed.

MEANS OF GDS VARIABLES: STANDARDIZATION SAMPLE

	6-7 year olds				8-9 year olds				10-11 year olds			
	Boys		Girls		Boys		Girls		Boys		Girls	
Delay Task (n = 192)	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Total Rewards	40.0	9.0	36.6	12.0	42.8	10.5	42.2	10.4	47.3	10.5	45.9	8.4
Total Responses	52.6	15.2	47.1	17.7	53.2	14.9	52.0	14.2	54.4	13.7	53.8	9.8
Total Efficiency Ratio	.80	.16	.83	.18	.83	.13	.83	.13	.88	.09	.86	.11
Vigilance Task (n = 91)	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Total Correct	30.7	9.4	25.7	11.8	37.6	6.0	38.5	3.7	41.8	1.7	41.5	2.3
Total Omissions	14.3	9.4	19.3	11.8	7.4	6.0	6.5	3.7	3.2	1.7	3.5	2.3
Total Commissions	16.1	15.5	9.3	5.0	11.8	10.7	9.6	10.3	6.9	8.0	5.1	4.1

Summary

While the GDS has successfully differentiated between diagnostic groups, the results are often difficult if not impossible to interpret because of the absence of reliable, independent criteria for making the classification of hyperactivity. Even when multiple criteria are employed the degree of diagnostic agreement among the criterion measures tends to be poor. Consequently, one can never be at all confident that misclassifications by the GDS according to other criteria are really errors of classification at all or are rather representative of more accurate assessment. Our experience to date has led us to gear future research away from criterion-based studies to those which focus on behavioral correlates of GDS performance.

264

CLINICAL APPLICATION OF THE GDS

INTRODUCTION

Effective utilization of the GDS rests upon the practitioner's understanding of the various processes involved when a child performs the tasks. The clinician must take into account the demands that the tasks place upon a child, the significance of the scores generated, and the value of the qualitative observations afforded during the performance of the tasks. The actual scores provide a basis for objectively comparing a child to a large group of non-ADD children. However, they must be interpreted with an understanding of the attributes and behaviors that affect a child's performance. Integrated with other observations, they can lead to making a diagnosis or classification and guide the practitioner towards prescribing an environment that is appropriately designed to address the individual child's deficits and strengths.

In the following analysis, we will be dealing with abstract concepts such as "arousal", "attention", "mediation", and cognitive "processes". These are difficult to define and even more difficult to study objectively. We are presenting a large body of objective data derived from our experience testing a wide range of children with the GDS. However, much of what follows is the conceptual construct we have developed out of this experience.

THE GDS DELAY TASK

Demands Of The Task

The Delay Task has been demonstrated to measure a child's ability to suppress or delay impulsive behavioral responses. In spite of its apparent simplicity (see Description of the GDS on page 3), this task embodies a variety of processes which are quite complex. For this discussion, we have divided them into three groups.

1. Cognitive Skills
2. Behavioral Demands
3. Motivational Factors

Cognitive Skills

One of the major benefits of the GDS Delay Task is that a child's performance is relatively independent of measured IQ (See page 5). Nevertheless, a host of perceptual and cognitive skills are exercised in the performance of this task and need to be considered in the interpretation of the results.

Consider the situation which confronts the subject. He is instructed that, to win a point, he must "push the blue button, wait a while, then press it again", but that if the button is pressed prematurely, no reward will be obtained. These instructions are simple enough to be understood by most six-year olds and even some younger children. The reward is the illumination of a light and an incrementing number, both easily observed by a child of any age, even one with a moderate visual impairment.

As the duration of the interval is not mentioned in the instructions, the child must develop a method of determining the minimum time he must refrain from responding to obtain a point. Many non-impulsive children will find this interval through a series of successive approximations. That is, they will wait a relatively long period, press the button, be rewarded for waiting long enough, and then progressively shorten the duration of their delay time until they arrive at the approximate minimum inter-response time which will usually result in a reward. Some non-impulsive children will take the opposite strategy and gradually increase the time between responses until they find a successful interval. Other children will use a combination of both approaches to maximize their reward.

While this process of "fine tuning" responses seems simple, it actually constitutes a very sophisticated approach. The child must be able to efficiently utilize the feedback provided to guide his responses. He must also develop a strategy for estimating each interresponse interval throughout the remainder of the task.

Behavioral Demands

In addition to the cognitive ability required, successful performance of the task requires certain behavioral skills. While the child is determining the necessary interresponse interval, he must refrain from responding until the intended interval has elapsed. Since the reward is only obtained when the button is pressed, the self-restraint can be difficult. Once the initial strategy is devised, the child must maintain the suppression or delay of his responses for the appropriate interval throughout the remainder of the task.

The strategy used to repeatedly wait the appropriate interval involves a behavioral skill. Most non-impulsive children use a "covert" strategy, some form of internal, cognitively-oriented delaying procedure, usually counting silently. Our research has indicated that this is the most efficient method, in that children who employ this strategy are usually quite adept at refraining from impulsive responses. While the notion of counting to one's self may seem simple to an adult, it is actually a very sophisticated cognitive strategy for a child to devise independently. Our experience has shown that children who employ more "overt" or behaviorally manifested strategies tend to perform much less efficiently on this task.

Motivational Factors

While children are almost universally captivated by the task, the apparatus itself is designed to be simple, non-distracting and unrewarding. The motivation to perform well reflects both the child's overall motivation to "do well" and the child's receptivity to

265

the task's rewards. Since the task demands the application of cognitive and behavioral skills, the child must be reasonably motivated to maintain a good performance throughout the task.

The subject is initially told that the task is a game in which he will have the chance to win a lot of points. This helps to instill in the subject a motivation to meet the challenge of "beating" the machine. A tone sounds to begin the game and sounds again at the end of the game along with the illumination of a green light. The brief illumination of the red light and the incremental counter function as the reward or "feedback" to provide the subject with reinforcement for a correct response.

The desire to gain control over a situation and the willingness to persevere with a rote task are important factors in successful coping. In fact, a respected theory of hyperactivity construes the problem as a developmental disorder of motivation (Glow & Glow, 1979). The notion is that these children are relatively unresponsive to feedback from the environment and come to see events in the world as generally unconnected to their own efforts. As a result, they are less aware of cause and effect relationships and fail to develop a sense of competence and intrinsic motivation. The motivation to succeed and the desire to master the situation are very key components of performance on the Delay Task. Without the motivation to bring resources to bear, adequate coping is impossible.

In summary, there are a variety of processes which enter into efficiently performing the Delay Task. It requires the cognitive ability to devise and execute a strategy for waiting long enough to get rewarded. The child must have the behavioral control to suppress or delay the impulses to press the button too early. To achieve both of these, a substantial degree of motivation is needed. All of these processes must function in concert for the child to perform well. It is important to note that, even though the Delay Task is relatively long and mundane, it is not a measure of the ability to sustain attention, **per se.** Constant attention is not required and the child's interest is repeatedly renewed by the reward. In fact, the demands that the Delay Task places upon the attentional processes are minimal. A child could daydream, play, or be otherwise distracted during the Delay Task and still perform adequately on the same measure. He might even use those activities as strategies to achieve the necessary interresponse interval. Focusing and sustaining attention usually facilitate Delay Task performance, but are not essential. Rather, the Delay Task places maximum demands upon the ability to delay or to refrain from emitting non-reinforced responses.

Data From The GDS Delay Task

The GDS Delay Task generates a plethora of both quantitative and qualitative data about a child's performance. Although they have been described in previous sections, we have outlined below the major categories of information yielded by the task:

1. Total scores
   a. Rewards--Number of times response button pressed after Interval has elapsed (correct "hits")
   b. Responses--Total number of times button was pressed
   c. Efficiency Ratio = Rewards/Responses
2. Block Data - Rewards, Responses and Efficiency Ratio for each of 4 blocks
3. Task Monitoring Data (available using the GDS Data Analysis Program)
   Distribution of interresponse intervals
4. Subjective Data
   a. Strategy - "cognitive" vs. "physical"
   b. Bursts of responses
   c. Other unique responses

Interpretation Of The Delay Task

Interpreting the results of the delay task requires the practitioner to evaluate both the objective and subjective data. Evaluation of the objective data is facilitated by comparing results to the normative data (see page 5) and to results we have obtained testing subjects in clinical settings (see page 4). We have found that in normal children, the performance does not differ significantly among the blocks; there is no significant learning effect or deterioration with time. We therefore feel that you should first evaluate the total scores and only need to examine the block scores closely if the total scores suggest abnormal responses.

Evaluation of the subjective data is less definitive, but in the discussion that follows, we have presented an overview of our experience testing a variety of children. While the subjective information is invaluable for assessing the child and planning appropriate learning environments, your interpretation will be strongly affected by your professional training.

Evaluating Total Scores

In the allotted eight minutes of the delay task (Standard Parameters), there is the theoretical opportunity for 80 Rewards. Since there is an initial adjustment period and no one can time their interresponse interval precisely, a score of 80 points is unrealistic. Our experience has shown that most non-impulsive children obtain between 30 and 50 points out of 35 to 65 total responses, giving an efficiency ratio of .70 to .98.

The Efficiency Ratio has been demonstrated to be the best

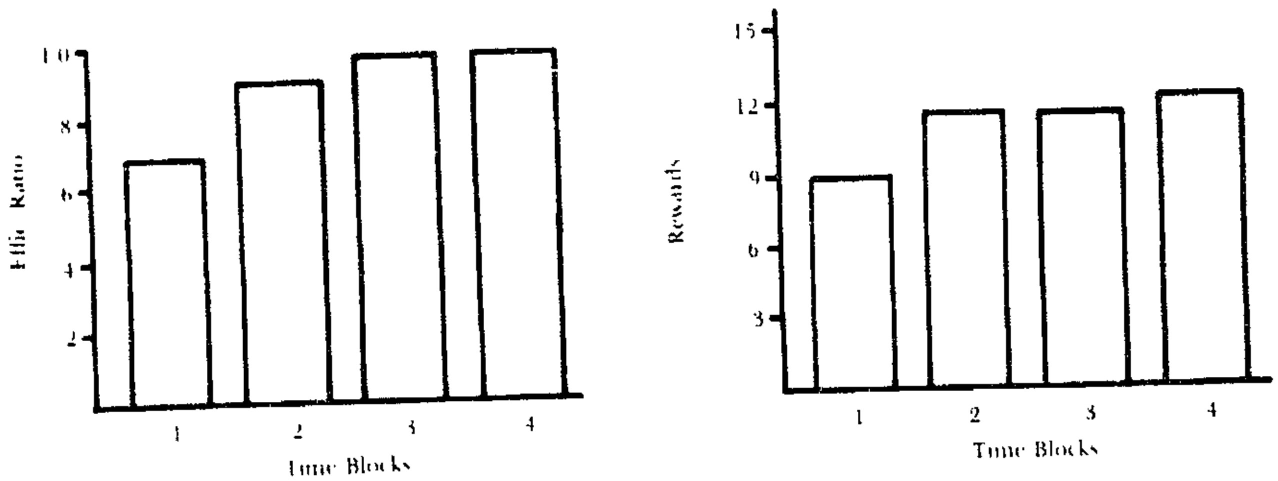
Delay Task index (see page 4) for discriminating between normal children and those judged to be hyperactive by a variety of other criteria. A child who cannot inhibit the frequent impulses to press the button will have a high number of total Responses. He will also tend to gain fewer total Rewards, because most of his responses will have occurred too soon after a previous impulsive response to be rewarded. As a ratio of Rewards to Responses, the Efficiency Ratio best reflects the effects of impulsivity on task performance.

However, the Efficiency Ratio can be misleading and needs to be interpreted in conjunction with the absolute number of Rewards. A child who presses the button only once every two minutes will have an Efficiency Ratio of 1.0, but is clearly different from the normal children we have tested. If a child has a normal Efficiency Ratio (> 0.55), but too few Rewards (< 18) several possibilites are evident. Most frequently, these scores are interpreted in terms of motivational factors such as resistance, acute stress or rebelliousness. Sometimes a child may recognize that he lacks the patterns of control necessary to perform well and overcompensate by ignoring the task for long periods.

Evaluating Block Scores

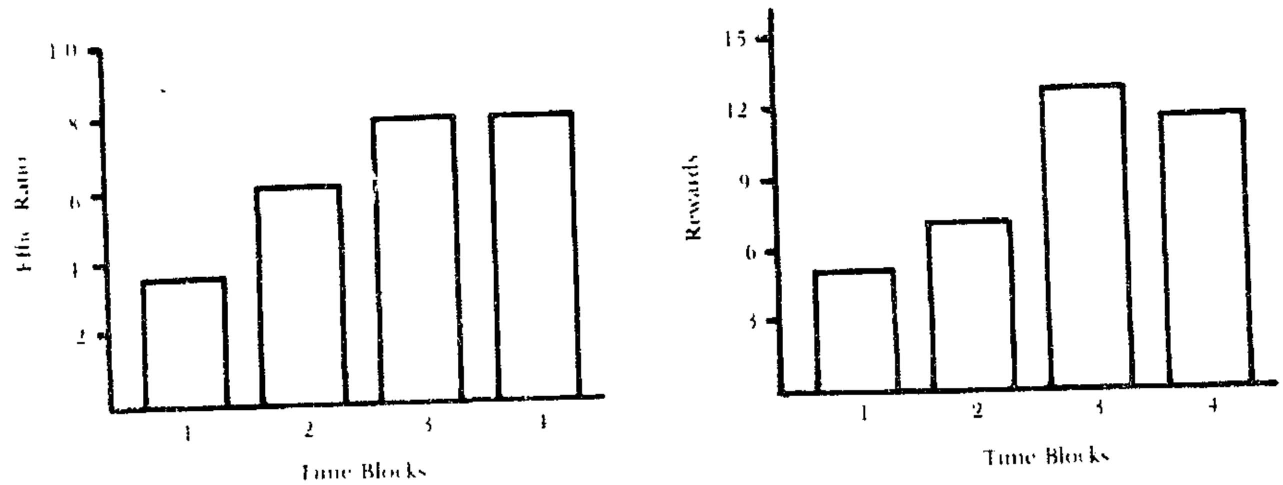
When the Total Efficiency Ratio or Total Rewards scores are abnormally low, more information about the subject can often be developed by examining the scores obtained in each of the four time blocks of the task. Examples of block by block analyses are presented below.

Profile 1



This is the profile of a youngster who demonstrates a normal overall performance on the task. We have found that such children do not significantly vary their performance from one block to another. The presence of impulsive behavior is not substantiated by these results and concern about behavior in such a child should focus on factors such as environmental stress, anxiety, or family issues.

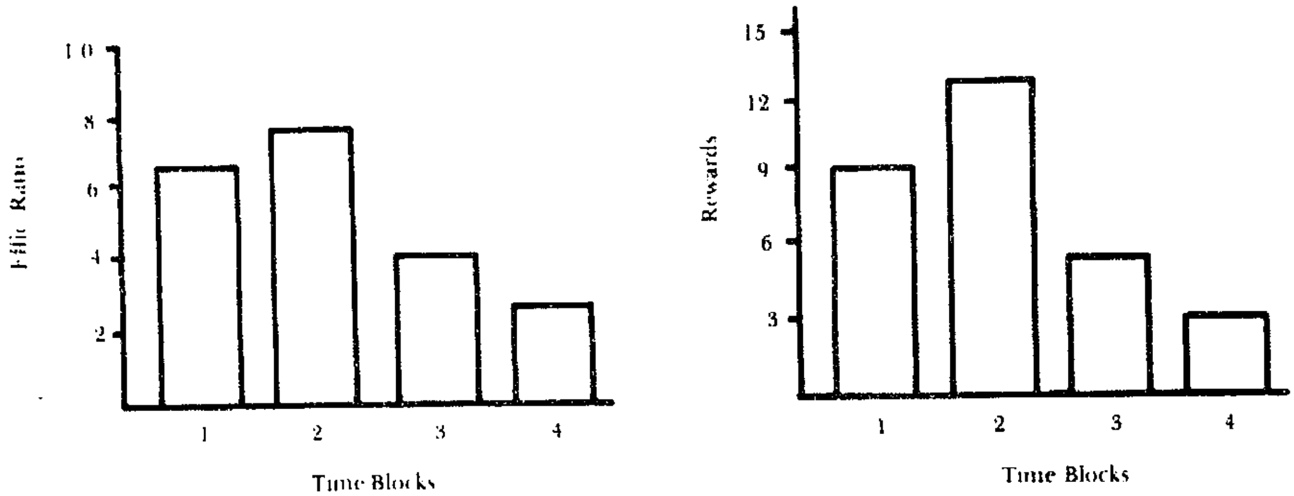
Profile 2



This profile is that of a youngster who has a difficult time initially adjusting to the demands of the task, but who can ultimately "adapt" and develop the strategies necessary to perform adequately. The relatively normal number of Rewards in the early Blocks indicates that the excessive Responses occurred in bursts that were widely enough separated to allow points to be earned. Later in the task, the child was able to suppress these bursts. Concerns about impulsivity should be addressed by emphasizing structure and considerable feedback in the early stages of any task the child is learning to perform.

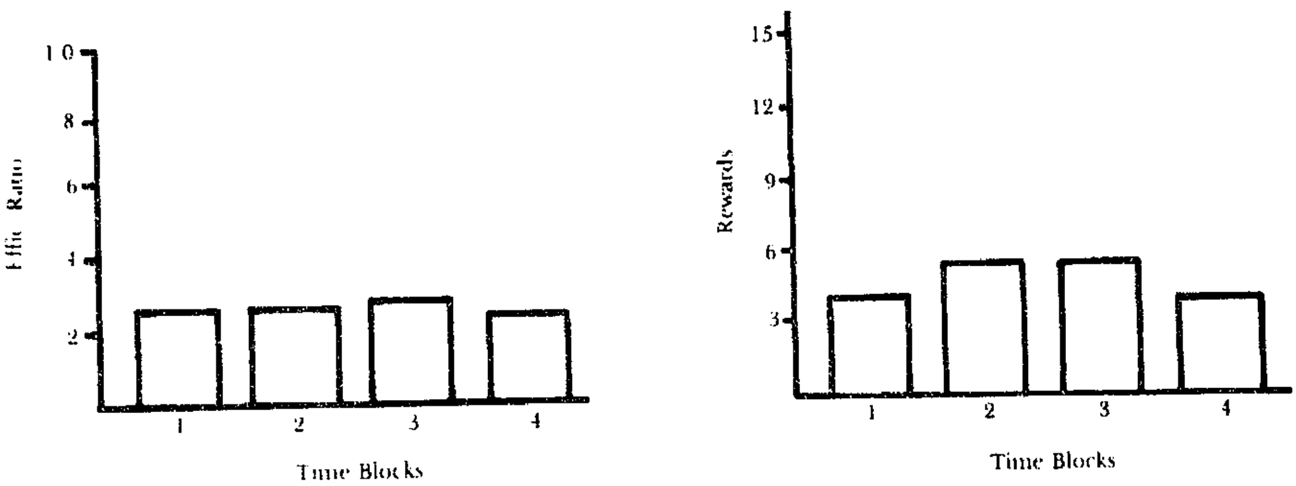
267

Profile 3



This profile is suggestive of a child who is capable of delay and suppression, but who has difficulty sustaining motivation or attention. In this case, the Rewards and Responses both diminish with time, indicating a general decrease in the child's participation in the task. Since the child is initially capable of nonhyperactive performance, issues such as environmental stress and social history may provide important information to help with the behavioral concerns

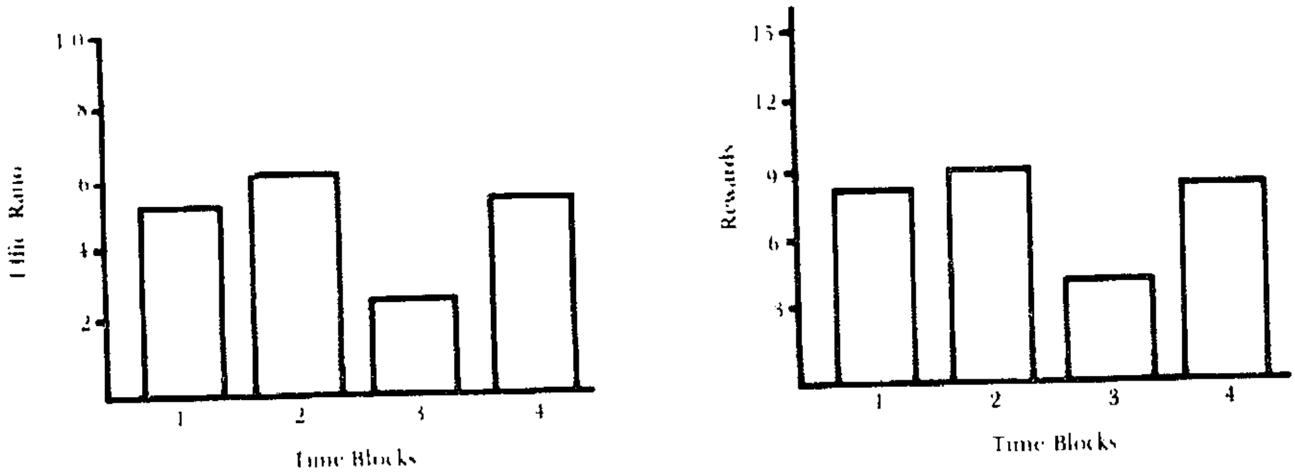
Profile 4



This pattern reflects a youngster with significant deficits in the ability to inhibit his impulses to press the button. As described under TOTAL SCORES, this behavior diminishes both the Efficiency Ratio and the Rewards. The uniform results across the time Blocks indicate that the child never mastered the necessary skills to perform well. The suspicion of impulsive behaviors is confirmed by this task performance pattern. Interventions should be geared toward cognitive-behavioral strategies and possibly stimulant medication.

268

Profile 5



This pattern is rather unusual, but has appeared from time to time in our clinical practice. It reflects a marginal ability to suppress responding with a rather profound lapse in the third time Block. While no definitive interpretations are offered, it is interesting to note that most children demonstrating this pattern were referred for emotional problems rather than assessment of impulsivity. This relationship will be investigated further in the future.

Evaluating Task Monitoring Data

When used in conjunction with the Data Analysis Program, the GDS provides data about the details of a child's performance during the Delay Task. The computer can display the distribution of interresponse intervals for each of the four time Blocks and for the total task. This data should be a valuable adjunct in determining why a child did poorly on the task. The child who has a low Efficiency Ratio from too many bursts of button presses will have a high frequency of response intervals in the 0-2 second range. In contrast, a child who is less impulsive may have a poor Efficiency Ratio with a high frequency of "near misses", interresponse intervals of 4-5 seconds. Research is being performed to establish interpretive standards for evaluating the interresponse distribution, but the practitioner may already be able to use the data to confirm subjective impressions about the child's difficulty with the task.

Subjective Data

As with any psychometric technique, observations of the child's behavior can be as telling as the scores themselves. Children without ADD tend to sit quietly and exhibit little actual movement throughout the session. When asked what they did to help them wait, the majority of children over 7 years of age say that they counted and that the game was easy. Younger children may display somewhat more activity than older children, but they are generally well-controlled and organized in their approach to the task. Children with an impulsive style, on the other hand, tend to have profound difficulty sitting and maintaining their motivation. They will squirm, leave their seat, run around the table, talk to the examiner and engage in a variety of other behaviors between responses. Many of these children will attempt to manipulate the apparatus by pulling at the button, attempting to open the back panel, or by pressing on the stimulus lights. (One particularly active boy tried to throw the device down a flight of stairs.) Again, it is rare for a child without an attention deficit to exhibit such behaviors. The tasks are as uneventful for them as they are trying for there ADD counterparts.

There are behavioral patterns which are characteristic of certain groups of non-ADD children. As one might expect, unusually fearful children will tend to sit extremely still and produce very few responses. One highly anxious child from an outpatient clinic pressed the button only once in eight minutes and sat frozen for the duration. Very competitive children can become excited and challenged but will still achieve normal scores. Observations of the child's behavior during the tasks yield a bounty of information and should be recorded carefully.

GDS VIGILANCE TASK

Demands Of The Task

The GDS Vigilance Task yields data regarding the child's ability to focus attention on a task and to maintain this attention over time in the absence of reinforcement (see Description of the Task on page 4). While the primary demand of this task is maintaining attention, a child's successful performance also entails other skills as outlined below.

1. Cognitive Skills
2. Maintaining Attention
3. Behavioral Demands
4. Motivational Factors

Many of the comments already made in reference to the Delay Task also apply to the Vigilance Task. They have not been reiterated in this section so the reader is strongly advised to read the discussion of the Delay Task (pages 6 to 10).

269

Cognitive Skills

The Vigilance Task and its instructions are quite different from the Delay Task, but the cognitive skill requred to understand the instruction set is similar. A Correct response is registered only when the subject presses the button after a "hot 9", that is a "9" immediately preceded by a "1". These instructions require the subject to attend to two stimuli:

a. the "1" serves as the "alerting stimulus" which prepares the subject to respond

b. the "9" serves as the "target stimulus" to which the subject must respond only after having been alerted.

Unlike the Delay Task, the Vigilance Task required modification for younger children. Those under 6 years may have difficulty understanding that they are to push the button only when a "9" is immediately preceded by a "1" on the digital display. Some do not even recognize the number nine yet. The "1" mode of the Vigilance Task (page 4) is simple enough for these children. (We have found that young children perform best with a Presentation Interval of 2 seconds and a Block Lengths of 120 seconds).

Maintaining Attention

The Vigilance Task places its maximum demand upon the child's ability to maintain attention. This demand is exacting because the Task requires constant attention for nine minutes (Standard Parameters) and because no reinforcement is provided to the subject during that time. The subject's responses must be guided by his attentional and cognitive processes alone. Children who are hyperactive usually do poorly on the Vigilance Task because they cannot resist the impulse to press the button at inappropriate times, or because they miss the stimuli when distracted. Children who are not hyperactive may also do poorly on the Vigilance Task because they cannot sustain the necessary attention to press the button at appropriate times.

Behavioral Demands

In order to do well, the subject must not only respond to every "hot" target stimulus, but he must also suppress or refrain from the impulse to respond to the alerting stimulus, non-target numbers which immediately follow the alerting stimulus, and a "non-hot" target stimulus. Highly impulsive children must also resist their tendency to respond randomly. Therefore, impulse suppression is an important demand in the Vigilance Task, although the primary demands are in the area of sustaining attention.

Motivational Factors

The general motivation to succeed on this task is essentially the same as that required on the Delay Task. However, a very critical difference between the two tasks is that here the subject must pay attention for nine minutes in the absence of any external reinforcement. The apparatus provides no feedback whatsoever; the subject receives no information regarding his performance. The only feedback available is that the subject will often be aware of his errors immediately after committing them. Linked with the general motivation to do well, this internal feedback can help to sustain his attention.

Data From The GDS Vigilance Task

Outlined below are the various scores generated by the Vigilance Task:

1. Total scores
   a. Correct--responses to a "1/9" sequence
   b. Omissions--failure to respond to the "1/9" sequence
   c. Commissions--responses to other than the "1/9" sequence
2. Block Data - Correct, Omissions and Commisions for each of 3 blocks
3. Task Monitoring Data (available using the GDS Data Analysis Program) - Distribution of digit patterns preceding the responses
4. Subjective Data

Interpretation Of The Vigilance Task

The general comments made for the Delay Task (page 7) apply equally well to this task. The practitioner should evaluate all the available information before making a diagnosis and planning a therapeutic approach.

As mentioned earlier, the Vigilance Task is a version of the Continuous Performance Test developed by Rosvold and his colleagues in 1956. Although the technique has been employed in many research studies through the years, formal norms were not established until standardization studies for the GDS were completed. For technical reasons, the Vigilance Task could not be included until these studies were well underway, and our clinical experience with the task is not as extensive as that for the Delay Task. A complicating factor is that Vigilance Task scores, particularly errors of Omission, are highly age-related (see page 5), and there is a great deal of variability under the age of 8 years. Consequently, the thresholds for interpretation described below should be used with care, especially for the records of younger children. Research is in progress and should considerably increase the diagnostic power of the task.

The Vigilance Task yields two fundamental sets of results: errors of Omission, and errors of Commission. Omissions are registered when the "1/9" sequence occurs, but the subject does not respond. The alerting stimulus, the "1", did not have its intended function of preparing the child to respond to the target stimulus, "9". These errors are most frequently due to inattention, but can arise from failure to understand the instructions or insufficient motivation. (Since the total number of occurrences of the "1/9" sequence is determined only by the task Parameters, the number of Correct responses is a measure that is dependent on the number of Omissions. The number Correct is the number of "1/9" presentations, 45 in 9 minutes, minus the Omissions).

The second independent measure, Commissions, is the number of responses to any digits other than the "hot" target stimulus, i.e. a response which occurred after a number sequence other than a "9" immediately preceded by a "1". Errors of Commission fall into three basic categories:

a. redundant errors,

b. "impulsive" errors, and

c. "random" errors.

While all three types are considered to connote impulsivity, they embody different nuances of impulsive behavior

Redundant errors are those which occur when the subject responds to the "hot 9" correctly, and then responds to the very next digit that is presented. This is considered to be a "motor overflow" error, and indicates that the subject is unable to refrain from emitting more than one response once he has been "primed" by the hot sequence.

Impulsive errors reflect any response made to the alerting or target stimulus when it is not appropriately accompanied by the other. A response to the alerting stimulus suggests that once the subject has been "primed", he is unable to suppress the response long enough to wait for the presentation of the target stimulus. A response to a "non-hot 9" similarly reflects an inability to refrain from responding.

Random errors are responses which occur to numbers other than "1", "9", or any number immediately following the "1/9" sequence. These errors are mostly likely linked to lapses in motivation or misinterpreting instructions. (Data for categorizing the errors of Commission is made available by the Data Analysis Program).

Four basic Patterns emerge in Vigilance Task performance:

1. Omissions: Normal*(< 15)
   Commissions: Normal (< 28)

This pattern reflects a subject who has sustained the attentional processes sufficiently to "catch" the expected number of "1-9" sequences, and who can also refrain from responding to non-target stimuli. These youngsters are also not impulsive

2. Omissions: Abnormal (> 15)
   Commissions: Normal (< 28)

This pattern suggests a child whose major difficulties lies in the area of sustained attention or distractibility. There is little to suggest difficulties in the area of impulse suppression. Many of these youngsters are classified as learning disabled or as ADD without hyperactivity.

----

*Thresholds are approximate and were determined for children over 8 years of age.

3. Omissions: Normal (< 15)
   Commissions: Abnormal (> 28)

This pattern reflects a subject who can attend sufficiently in a sustained fashion, but who cannot refrain from emitting impulsive responses.

4. Omissions: Abnormal (> 15)
   Commissions. Abnormal (> 28)

This pattern is reflective of a child who not only is highly distractible and inattentive, but who is also quite impulsive. Children who present with this pattern of responses on the Vigilance Task may have some form of learning disability in addition to impulsivity which gives rise to hyperactive behaviors. They will also tend to do poorly on the Delay Task.

In summary, these two tasks provide the clinician with a wealth of information regarding a child's abilities to suppress impulsive responses and to focus and sustain attention. Our clinical work using the two tasks in tandem has been very successful, primarily because together they cover a broad range of abilities that are essential to academic performance and behavioral stability. Employing these two tasks together has enabled us to understand some situations which otherwise would have exceeded our abilities to conceptualize and treat a particular child's difficulties.

TASK INTERRELATIONSHIPS

Our research has indicated that the two GDS tasks tend to correlate in the .58 - .60 range. Children who do poorly on the Delay Task also tend to do poorly on the errors of Commission of the Vigilance Task. By the same token, children who perform well on the Delay Task tend to make few errors of Commission. Errors of Omission, on the other hand, tend to operated fairly independently of the measures geared toward the assessment of impulsivity. Thus, while the two measures tend to correlate at a relatively high level and apparently tap into some similar dimensions, in a variety of ways they also measure different abilities.

One major difference between the two task has to do with the nature of reinforcement or feedback provided by the apparatus. The provisions of feedback is an integral function of the apparatus in the Delay Mode. In the Vigilance mode the apparatus provides no feedback. The child must perform for nine minutes essentially ignorant of the success or failure of his responses. This helps to make the Vigilance Task a pure measure of attention.

From a theoretical perspective, then, it may be that the dimension of informational feedback may account, in part, for a child's performance difference between the two tasks. It could be hypothesized that children who perform well on the Delay Task and poorly on the Vigilance Task are in some ways communicating something to us regarding their ability to perform in the presence or absence of informational feedback.

CLINICAL EXAMPLES

Case 1

A seven year old girl was referred to the child development clinic by her teacher for poor school performance, not following instructions, missing assignments, and difficulty with reading. The teacher viewed much of this youngster's behavior as willful and oppositional, and was at her wit's end as to how to help the girl learn. Her parents reiterated the teachers complaints, and added that they felt the child to be distant and aloof much of the time. During the clinical interview this youngster was pleasant and cooperative, but gave the impression of either wanting to be somewhere else or of simply daydreaming.

On the Delay Task she obtained an overall Efficiency Ratio of 82, well within the normal range. She earned 46 Total rewards. Her responses appeared controlled, orderly, and goal directed. Obviously her difficulties did not lie in the area of impulse control. On the Vigilance Task her difficulties became quite clear. Although she produced just one error of Commission, which is well within normal limits, she made 33 errors of Omission, scoring well beyond the mean on this measure.

Although her motivation throughout the task was quite good, she was unable to maintain the degree of preparedness required to perform effectively on this part of the task. Therefore, we were able to identify her main difficulty as a deficit in sustained attention in the absence of impulsivity, which fits the DSM III classification of Attention Deficit Disorder without Hyperactivity.

She was placed on a very small dose of stimulant medication, and her teacher was informed of the results of the testing. Intervention strategies were offered for both home and school, and her adjustment improved in a satisfactory manner.

Case 2

The two GDS tasks used together may also be of considerable help to the clinician in identifying children who are **not** impulsive or appropriate to classify as ADD. A ten year old boy was referred to this same clinic for the evaluation of learning problems, difficulties with classmates, much "out-of-seat" behavior, and a variety of other behavior management problems. The parents shared the school's concerns, and were very eager to obtain some pharmaceutical cure for their son.

Three years prior to this evaluation, he had been diagnosed as hyperactive and had been placed on stimulant medication. His parents reported that his initial response to this regimen was good in that he became less active and less of a management problem. They reported that this "honeymoon" period was short-lived, and he quickly became grouchy, irritable, and somewhat depressed. Preferring to have the boy active and boisterous, they terminated his stimulant medication after consulting a pediatrician. However, his problematic behavior persisted, and by the time of this referral, had escalated to the point where some form of intervention was urgently indicated.

This youngster's performance on the Delay Task was solidly in the average range (.85) with 58 rewards (normal). His responses were orderly and controlled. While he appeared fidgety and restless during his task performance, he did quite well. On the Vigilance Task his scores were also within the normal range, with two errors of Omission and only one error of Commission. These results firmly suggested that the behaviors which were bringing him into conflict with his environment stemmed from factors other than attentional deficits and impulsivity. Although his behavior truly was impulsive from a descriptive standpoint, he did not demonstrate any deficit in the ability to suppress responding or sustain attention.

Following our initial evaluation, psychological testing was undertaken to identify the sources of his problem behaviors. It was found that this youngster was frequently overwhelmed by the high goals that he perceived others had for him. He was terrified of failing in school and felt he could never fulfill his parent's expectations. This gave rise to considerable anxiety, which he discharged motorically through overactivity and impulsive behavior. He was placed in outpatient psychotherapy with a family emphasis, and currently is reported to be doing quite well.

Case 3

A nine year old boy was referred with complaints of restlessness, noncomplicance, fighting with other children, and poor academic performance. While his parents had been recently divorced, these behaviors were longstanding and had been problematic since early childhood. A medical examination by his pediatrician was unremarkable, and he was referred for a psychological evaluation. Upon initial contact with the psychologist, this youngster appeared inhibited, quiet, and withdrawn, giving no indications of impulsive or hyperactive behavior. However, when confronted with the demands of the Delay Task, the boy was unable to maintain the illusion. He achieved an efficiency ratio of .44 (less than the 5th percentile) with 28 Rewards. His behavior throughout the Delay Task was disjointed, involving much out-of-seat behavior and extreme restlessness. While these behaviors in some ways helped him to suppress responding, because he was otherwise occupied, he was nonetheless unable to refrain from emitting a large number of unreinforced responses. This pattern of behavior was repeated during the Vigilance Task. He had a high number of "random errors" and consistently responded immediately after a "1". His performance was clearly suggestive of an inability to delay, particularly once he had been primed to respond. The youngster was placed on a moderate dose of stimulant medication, and was classified as having ADD with Hyperactivity. His academic programming was geared more toward accuracy than speed, and he received resource help to encourage him to modulate his response style. Follow-up contact indicated substantial improvement.

ONGOING RESEARCH

Our own research efforts concerning the clinical application of the GDS will include studies on effects of reinforcement and feedback, as well as research on the effects of stimulant drugs, patterns of responses of learning disabled children, and the possible applications of the GDS in training children to be less impulsive and more attentive. One of the major enhancements that has been developed for the GDS is an interface with a portable computer. In addition to making the total and block data more accessible, the computer has been programmed to directly receive GDS data while a task is being performed. For the Delay Task, the computer stores every interresponse interval and summarizes this data by generating a listing of the interresponse intervals. For the Vigilance Task, it saves the record of every button press and assembles a list of the frequency of digit sequences preceding the presses. This data will enable the clinician to make even more precise determinations of **why** the subject performed in certain ways.

For example, when a child makes numerous errors of Commission on the VigilanceTask, it will be very important to the clinician to know where these errors occurred in the sequence of the task. If they occur because the child tries to respond to the hot combination, but responds too late to be credited with a correct response, this is quite different from the child who responds to the "9" when it is not hot or who responds to the "1" alone. This computer innovation will have profound and exciting implications for our ability to "fine tune" our interpretations of the data provided by the GDS.

REFERENCES

Achenbach, T. The child behavior profile: I. Boys aged 6-11. *Journal of Consulting and Clinical Psychology,* 1978, 46, 478-488.

Achenbach, T. & Edelbrock, C. The classification of children's psychopathology: A reveiw and analysis of empirical efforts. *Psychological Bulletin,* 1978, 85, 1275-1301.

Buddenhagen, R. & Sickler, P. Hyperactivity: A forty-eight hour sample plus a note on etiology. *American Journal of Mental Deficiency,* 1969, 73, 580.

Cairns, E. & Cammock, T. Development of a more reliable version of the Matching Familiar Figures Test. *Developmental Psychology,* 1978, 14, 555-560.

Cohen, N., Douglas, V., & Morganstern, G. The effects of methylphenidate on attentive behavior and autonomic activity in hyperactive children. *Psychopharmacologia,* 1972, 22, 282, 294.

Conners, C. A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry,* 1969, 126, 884-888.

Douglas, V. Stop, look, and listen: The problem of sustained attention and impulse control in hyperactive children. *Canadian Journal of Behavioral Sciences,* 1972, 4, 259-282.

Egeland, B. & Weinberg, R. The Matching Familiar Figures Test: A look at its psychometric credibility. *Child Development,* 1976, 47, 483-491.

Gardner, R. *The Objective Diagnosis of Minimal Brain Dysfunction.* Cresskill, N.J.: Creative Therapeutics, 1979.

Gittis, A. & Gordon, M. Developmental analysis of behavioral dysfunction in rats with septal lesions. *Journal of Comparative and Physiological Psychology,* 1977, 94-106.

Glow, P. H. & Glow, R. A. Hyperkinetic impulse disorder: A developmental defect of motivation. *Genetic Psychology Monograph,* 1979, No. 100, 159-231.

Gordon, M. The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive boys. *Journal of Abnormal Psychology,* 1979, 7, 317-326.

Kagan, J. Reflectivity-impulsivity: The generality and dynamics of cognitive tempo. *Journal of Abnormal Psychology,* 1966, 71, 17-24.

Kendall, P & Wilcox, L. Self-control in children: Development of a rating scale. *Journal of Consulting and Clinical Psychology,* 1979, 7, 317-326.

Maynard, R. Omaha pupils given behavior drugs. *Washington Post,* June 1970.

McClure, D. F. & Gordon M. The performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. Paper submitted to the *Journal of Abnormal Child Psychology,* 1983.

Machover, K. *Personality Projection in the Drawing of the Human Figure.* Springfield, Ill., Charles G. Thomas, 1949.

Miller, R. G., Palkes, H. S. & Stewart, M. A. Hyperactive children in suburban elementary schools, *Child Psychiatry and Human Development.* 1973, 4, 121-127.

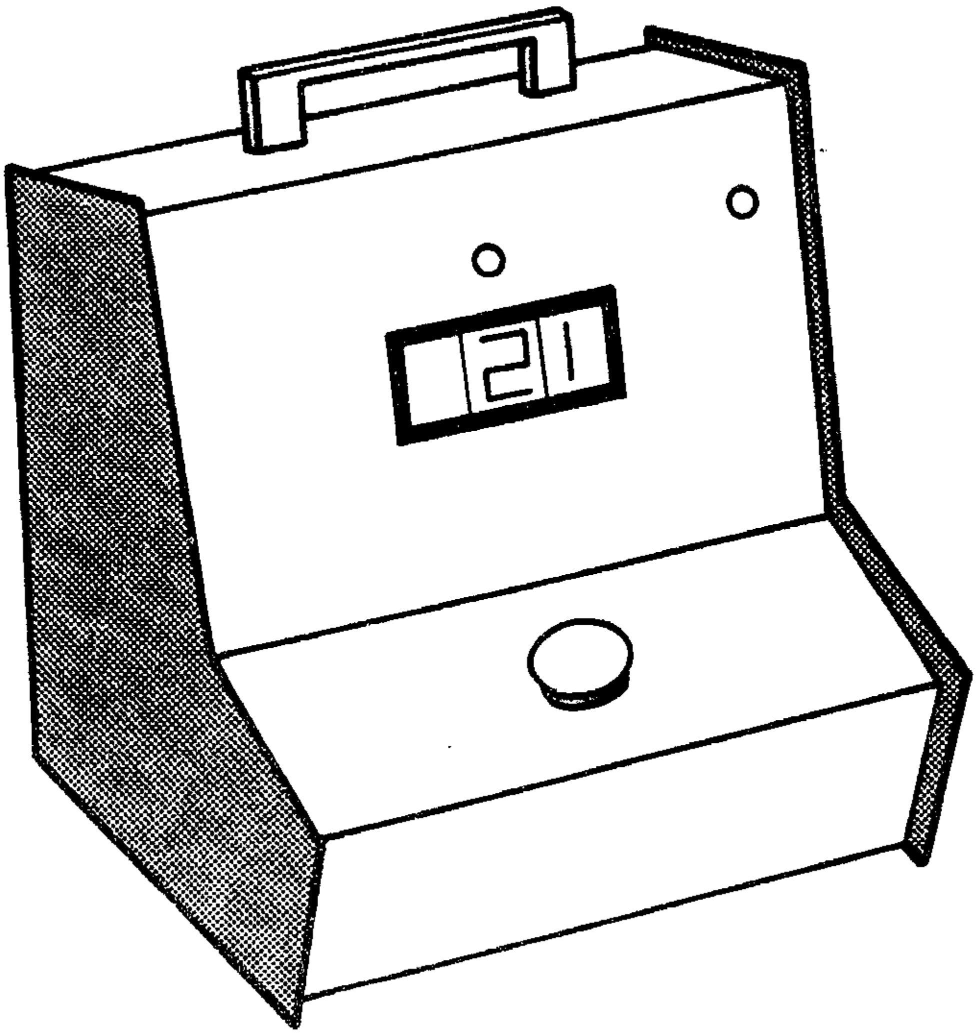
Rosvold, H., Mirsky, A., Sarason, I., Bransone, E., & Beck. A continuous performance test of brain damage. *Journal of Consulting Psychology,* 1956, 20, 343-352.

Stone, F., Wilson, M., Spence, M., & Gibson, R. A survey of elementary school children's behavior problems. Paper presented at the annual meeting of the American Orthopsychiatric Association, New York, 1969.

273

THE GORDON DIAGNOSTIC SYSTEM

INTERPRETIVE SUPPLEMENT



CLINICAL DIAGNOSTICS INC
1536 Cole Blvd., Golden, Colorado 80401
Suite 350 • 800-521-4503

274

SUPPLEMENT TO THE GDS INTERPRETIVE GUIDE

Michael Gordon, Ph.D.
Upstate Medical Center, Syracuse, New York

F. Daniel McClure, Ph.D.
D.C. Wilson Hospital, Charlottesville, Virginia

ACKNOWLEDGEMENTS

This Interpretive Supplement reflects the efforts of a talented and remarkably masochistic group. In particular, we would like to thank Ms. Barbara Bilinski Mettelman for her contributions both in coordinating our research and in managing the details of statistical analyses and other daily activities. The manuscript benefited greatly from the comments of Ms. Barbara Bass, Ms. Mary Galligan and Dr. Ernest Post from Clinical Diagnostics. Finally, we acknowledge once again the efforts of Dr. Wendy Gordon whose clinical and editorial expertise immensely enhanced this Supplement.

January 20, 1984

TABLE OF CONTENTS

INTRODUCTION

The purpose of this supplement is to provide specific guidelines for interpreting data generated by the GDS. The GDS taps into some very fundamental dimensions of functioning that are interrelated and affected by a host of other variables. The goal of the supplement is to assist the reader in gaining a thorough understanding of the Delay and Vigilance Tasks by providing a general framework for the interpretation and integration of the diverse factors inherent in GDS performance.

Some of the material from the original Interpretive Guide is repeated here for purposes of clarity and elaboration. We also want to point out that the logic underlying this Supplement served as the basis for the Data Analysis Program (DAP) we developed for use with microcomputers. The DAP generates the kinds of interpretations described in these pages but at a much faster rate.

As with other psychometric data, information from the GDS cannot be viewed in isolation; rather, it must be considered within the overall context of data from clinical observation, social/developmental history, and other psychodiagnostic procedures. In addition, individual GDS scores should only be interpreted in relationship to the other GDS scores and to the subject's overall performance on the tasks. The reader will find, however, that interpretation of GDS performance is straightforward and can be achieved with relative ease because it flows logically from a common-sense understanding of the tasks. It should be kept in mind that the goal of our interpretation is not only to establish normality or abnormality but also to understand how the child functions and translate that information into specific recommendations for treatment.

In this supplement we will first analyze the interpretive significance of individual scores from each test separately. Critical parameters will be identified and explained. We will then explore how the Task Monitoring data clarifies the data from each task. Finally, we will outline a method for integrating the data from both tasks into a sound and coherent overall interpretation of a subject's functioning. As we go through this process we will include interpretive statements that will help the reader understand how the authors describe and organize GDS data. These statements are not intended to be taken as cookbook interpretations. Rather, their purpose is to illustrate our thinking in a practical and concrete way.

Keep in mind that all interpretations are based upon administration of the tasks according to the Standard Parameters. It should also be noted that we do not discuss the "1" mode of the Vigilance Task. A separate supplement will cover the testing of very young children.

THRESHOLDS FOR INTERPRETATION

The first question to be answered for any GDS data is whether a particular score is normal or abnormal. Based upon data from standardization projects, we have divided scores into three ranges. A score is considered ABNORMAL if it is typical of less than 5 percent of the normal population, i.e., the 5th percentile or less. Scores in the BORDERLINE range are those which fall between the 6th and 25th percentile. Finally, scores that are above the 25th percentile are classified as NORMAL. These cutoffs correspond closely to clinical experience and our research studies in that children classified as hyperactive by other measures typically perform in the ABNORMAL range of GDS scores.

To facilitate the interpretive process, we have constructed Threshold Tables which list the ranges of scores corresponding to each category (ABNORMAL, BORDERLINE, and NORMAL). These tables are located at the end of the Supplement, and are reproduced in each section as needed. They are based on the same data as the standardization tables displayed at the end of the Interpretive Guide.

Because the Delay Task norms do not vary significantly with age, we have presented one table for all children from 6 to 12 years. Since some of the Vigilance Task variables are age-related, expanded tables were required. Comparison of an individual child's scores to these tables will determine the degree of normality or abnormality.

278

INTERPRETATION OF THE DELAY TASK

The GDS Delay Task yields three kinds of Summary Scores: the number of Responses (R); the number of Rewards or correct responses (REW); and the Efficiency Ratio (ER), which is a ratio obtained by dividing REW by R. For the purposes of interpretation, we pay most attention to the ER for the whole session – the Total ER.

The Efficiency Ratio (ER)

NORMAL	.77 – 1.0
BORDERLINE	.55 – .76
ABNORMAL	⩽.54

The ER is the basic building block of Delay Task interpretation. It represents the percentage of times the child pressed the button after having waited the appropriate number of seconds. The ER is the best single indicator of the level of impulsivity demonstrated by a subject. According to the Threshold Tables, an ER of .54 or less falls in the ABNORMAL range and indicates a significant deficit in the ability to inhibit responding. The chances are that a child with an ABNORMAL ER tends to display poor self-control and would be considered by most parents and teachers to be impulsive. Children with an ABNORMAL Total ER experience difficulty in situations which require them to come up with ways to keep themselves from engaging in activity. They are overly disinhibited and have not developed strategies to help them delay when delaying would serve them well. By the same token, scores between .55 and .76 are in the BORDERLINE range, and suggest tendencies toward impulsive responding. Obviously, the closer the ER approaches the lower end of the BORDERLINE range, the stronger the tendencies toward impulsivity.

Listed below are the kinds of interpretations and recommendations that can be formulated based on the Total ER:

Interpretive Statements

Total ER within the NORMAL range:

Subject demonstrates normal capacity for delay and inhibition. If this child were referred for impulsive behavior, a more extensive psychological evaluation should be considered because problem behaviors do not appear to stem from a fundamental deficit in impulse inhibition. Anxiety, emotional difficulties, and/or environmental stress may account for these behaviors.

Total ER within the BORDERLINE range:

Subject demonstrates tendencies toward impulsive responding. S/he will require increased external structure in the classroom, particularly when confronted with tasks which involve setting one's own pace and inhibiting behavior. This child might also benefit from special programs designed to teach strategies for developing cognitive self-control over behavior. In other words, this child will likely require help in learning how to "stop, look, listen, and think" before acting. Stimulant medication is not usually the treatment of choice for such a youngster, although in some cases a brief period of pharmacotherapy may facilitate other interventions.

Total ER within the ABNORMAL range:

Subject demonstrates significant deficits in areas related to the ability to delay. S/he will require greatly increased external structure in the classroom, particularly when confronted with tasks which involve self-control and the capacity to inhibit responding. This child would benefit from special programs designed to teach strategies for developing cognitive self-control over behavior. This child will require help in learning how to "stop, look, listen, and think" before acting. Treatment with stimulant medication may be a viable option to consider as part of a general program of educational and psychological management.

While the Total ER can provide an overall pegging of a child along the continuum of impulsivity, the fuller meaning of this score can be extracted by a careful analysis of the Block Scores and Task Monitoring Data (see below).

Total Number of Responses (R)

NORMAL	21 – 56
BORDERLINE	57 – 66
ABNORMAL	⩽20 & ⩾67

Total R reflects the subject's overall rate of responding. The primary significance of R is that it determines the validity of the Efficiency Ratio. A Total ER of .80 based on only 10 responses must be interpreted with caution because the child is not attempting to gain points at a rate typical of peers. Our experience indicates that children with low R are often fearful and overly-cautious. It is typical that further evalu-

279

ation shows them to experience emotional/motivational problems. It can also occur that very impulsive children refrain from responding at all for fear that any response would lead to a torrent of poor control.

Very high levels of R (especially 90 or more) also take on significance. These children exhibit a marked inability to suppress responding. They may be able to score Rewards but there is a flood of impulsive responding. Of course, most children with a high number of Responses will also have poor ER's and will be considered impulsive. When very high responsivity occurs, however, it is important to look for patterns of impulsivity (see below). In addition, the subject's motivation to perform well on the task should be considered suspect.

Total Number of Rewards (REW)

NORMAL	38 – 80
BORDERLINE	22 – 37
ABNORMAL	⩽ 21

Our data regarding the discriminative validity of the REW score has been equivocal. In one of our research studies, this variable significantly differentiated "hyperactive" from "nonhyperactive" groups. However, in another study REW failed to differentiate the experimental group from the control group. We tend not to interpret REW by itself, primarily because its impact is subsumed under interpretation of the ER.

The Analysis Of Block Scores

The GDS is internally programmed to separate the 8-minute trial into four 2-minute time blocks. This was engineered into the GDS because patterns of responding over time can be critical in the evaluation of attention and inhibition. Block Scores enable the practitioner to view how a given subject's performance has progressed over the course of the eight minutes and to detect changes in the pattern of responding. To accomplish this, the GDS yields values for R and REW for each two minute block. The resultant ER values are subscripted with the appropriate time block. For example the ER derived from Time Block 1 is notated as ER_1, the ER for Time Block 2, ER_2, and so on.

Data from both normal and clinical populations has demonstrated the GDS to have a remarkable degree of internal consistency, whether the subject is impulsive or non-impulsive. Performance is particularly consistent after the first block, during which time the subject adapts to the task. Therefore, marked inconsistency in a child's scores from block to block takes on clinical significance even for subjects who demonstrate the capacity to perform within the normal range on all variables.

To analyze Block Stores we consider 3 parameters: overall variability, slope, and extreme values.

Block Variability

NORMAL	0 – .10
BORDERLINE	.26 – .11
ABNORMAL	⩾ .27

Calculations for Block Variability are aimed at characterizing the overall degree of consistency across the session. For the purposes of statistical analysis and our Data Analysis Program, the standard deviation of the four Block ER scores is used as an indicator of variability. To compute the standard deviation precisely you would follow these steps:

1. Take the Mean (average) of the 4 ER Block Scores by adding them together and dividing by 4.
2. Subtract the Mean from each Block ER.
3. Square each difference and add the results of the 4 squares together.
4. Divide by 3 and then take the square root of the quotient.

To illustrate, a subject with Block ER's of .45, .67, .44 and .78 would have a mean ER of .59. Subtracting the mean from each ER and adding the square of each difference together yields a sum of .08. Dividing by 3 and taking the square root gives us a Block Variability score of .17. We can then compare that score to the Threshold Tables to judge overall consistency*.

A child with NORMAL Block Variability will be relatively consistent in his/her efforts to delay, regardless of whether those efforts are successful or not. Children with BORDERLINE or ABNORMAL scores will be less consistent.

Interpretive Statements

Block Variability NORMAL:

Subject demonstrates normal variability in performance, indicating that s/he will be no more inconsistent in delay or inhibition than peers.

*A simpler method for judging variability entails determining the Mean ER score (by adding the Block ER scores together and dividing by 4). If any Block ER falls above or below the Mean ER by more than .15 then there is a significant amount of variability. In the example given above there is significant variability in that one Block ER falls above or below the Mean ER of .59 by more than .15 (that is, the scores fall outside the bracket of .45 to .74).

280

Block Variability BORDERLINE:

Subject demonstrates a degree of variability which is in the borderline range. S/he is likely to be more inconsistent than peers in the level of impulsivity or inhibition demonstrated. Interventions should focus on the development of strategies geared toward achieving stability of responding. These might involve "talking" one's self through rote tasks, practicing step-by-step sequential tasks, and outlining on paper the various steps involved in solving a particular problem.

Block Variability ABNORMAL:

Subject demonstrates a significant degree of variability in his/her performance. S/he will be unpredictable in the level of impulsivity exhibited. As with those subjects whose variability is in the BORDERLINE range, interventions should be geared toward achieving and maintaining a stable, consistent response style. Such approaches might involve "talking" one's self through rote tasks, practicing step-by-step sequential tasks, and outlining on paper the various steps involved in solving a particular problem.

Slope

An important aspect of our interpretation of Block Scores centers around discerning patterns of variability. While the Block Variability score characterizes the general level of consistency, the Slope score determines if there were any clear trends from block to block. In other words, did the subject's scores progressively improve, stay the same, or deteriorate across the session? Again, most children's ER's are relatively stable from block to block. Inspection of the ER's, particularly when graphed, gives a sense of whether or not there was improvement or deterioration. For statistical purposes we use the following formula:

$$\text{Slope Score} = (ER_3 + ER_4)/2 - (ER_1 + ER_2)/2$$

If the Slope Score is between .12 and -.12, performance is characterized as having an Even Slope. Slope Scores of more than .12 reflect an Upward Slope (i.e., improvement) while scores below -.12 reflect a Downward Slope (i.e., Deterioration).

Interpretive Statements

Even Slope:

Subject demonstrates consistent performance across blocks.

Upward Slope:

Subject demonstrates initial tendencies toward impulsivity, but can adjust to the task and achieve normal responding. After the initial difficulty in adaptation, efficient performance is achieved. This child will require more structure and support in the beginning stages of a task or exercise than his/her peers.

Downward Slope:

Subject demonstrates a tendency toward deterioration in performance over time. May require additional structure to support responding as tasks progress temporally.

Analysis of extreme scores: Peaks and Valleys

Another important factor involves any Block Score that stands in contradistinction to the Total ER. For example, a subject with an ABNORMAL Total ER (below .55), but with at least one Block ER in the normal range (i.e., "a peak") is demonstrating some capacity for normal delay. In contrast is the subject with a similar ER but with no Block ER in the normal range (i.e., above .74). This subject at no point in the testing evidences a capacity to delay adequately. By the same token, a subject with a NORMAL Total ER, but with one Block ER below .55 (i.e., "a valley") is demonstrating a potential for impulsivity under certain conditions, while a subject with a normal Total ER and all normal Block ER's is consistently well-controlled.

Interpretive Statements

Total ER BORDERLINE OR ABNORMAL and any Block ER greater than .74:

This subject does have the capacity for normal delay.

Total ER BORDERLINE OR ABNORMAL and no Block ER greater than .74:

Performance never reaches the normal range.

Total ER NORMAL and 1 Block ER ABNORMAL (less than .54):

Despite overall normal performance, subject evidences some potential for impulsive responding.

281

Examples of Delay Task Interpretation

To clarify issues presented in the previous sections, let us look at some sample Delay Task patterns.

SUBJECT 1 – 8 year-old male

BLOCK	1	2	3	4	TOTAL
REWARDS	13	14	13	15	55
RESPONSES	17	16	16	17	66
EFFICIENCY RATIO	.77	.87	.80	.88	.84

INTERPRETATION: This subject's Number of Responses (R) indicates that these results may be interpreted with confidence, as responsivity is within normal limits. The Total ER indicates that he has no difficulties with delay. This ability to inhibit responding persists across time, suggesting that it is a relatively stable trait. This child does not appear to have an impulsive style, *per se.* We would suggest that the practitioner look toward other areas to account for the problem behaviors.

SUBJECT 2 – 10 year-old male

BLOCK	1	2	3	4	TOTAL
REWARDS	10	9	12	13	44
RESPONSES	27	19	23	22	91
EFFICIENCY RATIO	.37	.47	.52	.59	.48

INTERPRETATION: This subject's rate of responding (R) is abnormally high so that the results should be interpreted with caution. When the number of Responses is above 90, it indicates an extreme inability to suppress responding which can be related to motivational issues. A more thorough psychological evaluation would likely be suggested to evaluate the various components of the child's difficulties with inhibition.

This subject does indeed demonstrate significant impulsivity and inability to delay. While he improves in this area over time, at no time do his Block ER's approach the NORMAL range. He will require extra structure particularly at the initial stages of a task. He would also benefit from a training program aimed at developing strategies for self-control.

SUBJECT 3 – 9 year-old female

BLOCK	1	2	3	4	TOTAL
REWARDS	12	10	6	8	36
RESPONSES	15	17	15	19	66
EFFICIENCY RATIO	.80	.59	.40	.42	.54

INTERPRETATION: This subject's rate of responding indicates that these results may be interpreted with confidence, as responding is well within normal limits. She demonstrates a significant degree of impulsivity. However, she also demonstrates the ability to delay and inhibit in a normal manner in the initial stages of a task. This ability tends to deteriorate over time, suggesting that she will require additional structure as tasks progress.

Task Monitoring Data

Task Monitoring capability was programmed into the GDS in an effort to provide the user with a more detailed view of the subject's response style. Task Monitoring Data can only be collected with the GDS Data Analysis Program. While there are not yet enough normative data to make specific interpretive statements associated with Task Monitoring Data, they allow the user to "fine tune" interpretations of the other data generated by the Delay Task. The Task Monitoring Data provide a detailed summary of the response latency from the previous response for each response emitted. This data is provided for each 2-minute time block as well as for the entire test. An example is provided by the next subject.

SUBJECT 4 – 8 year-old boy

INITIAL RESPONSE LATENCY = 0.4 SEC.

INTERVAL	BLOCK 1	2	3	4	TOTAL
0 SEC	0	0	0	0	0
1 SEC	0	1	0	0	1
2 SEC	0	2	1	1	4
3 SEC	0	0	3	0	3
4 SEC	5	1	4	5	15
5 SEC	3	4	4	2	13
6 SEC	2	0	5	4	11
7 SEC	0	1	2	2	5
8 SEC	4	1	1	3	9
9 SEC	2	2	1	2	7
10 SEC	0	2	1	0	3
11 SEC	0	0	0	0	0
12 SEC	0	2	0	0	2
13 SEC	1	0	0	0	1
14 SEC	0	0	0	0	0
15 SEC	0	0	0	0	0
16 SEC	0	0	0	0	0
17 SEC	0	0	0	0	0
18 SEC	0	0	0	0	0
19 SEC	0	0	0	0	0
>19 SEC	0	0	0	0	0

In this case the Task Monitoring Data proved to be extremely helpful in analyzing data from the Delay Task. This youngster's other scores were as follows:

BLOCK	1	2	3	4	TOTAL
REWARDS	10*	8	10	11	39*
RESPONSES	18	16	22	19	75
EFFICIENCY RATIO	.55	.50	.45	.57	.52

*The first reward in Block 1 is given regardless of the initial latency and does not appear on the table of response intervals.

This protocol is valid (as indicated by R), and the subject has significant deficits in the ability to inhibit and delay responding (as indicated by ER). Moreover, the low degree of variability suggests that this trait is fairly stable, and Block ER's indicate that this subject was at no time able to achieve and maintain a normal level of responding. However, his Task Monitoring Data sheds a great deal of light on these scores. Analysis of this data indicates that 39 of his 75 responses occurred within 4-6 seconds of the previous response, while only 8 occurred within 0-3 seconds of the previous response. More than half of his responses clustered very closely to the 6-second criterion. This suggests that, although he did quite poorly on the test, he was indeed making some relatively stable attempts, albeit unsuccessful, to delay. His problem appears to be in arriving at a strategy that would help him delay the extra 1 - 2 seconds necessary to gain a point. For some reason, he did not undergo the process of "bracketing" or testing his responses so that they would be successful. It may be that this is a child who does not sufficiently incorporate feedback to correct and improve performance. In other words, something "got in the way" of his ability to use the unrewarded responses to adjust his strategy. We have found that some children who present with this protocol become so preoccupied and anxious about "winning" that they fail to settle down to a comfortable and productive session. For this reason, we often recommend that additional psychological testing be conducted to evaluate motivational/emotional factors.

This child's Task Monitoring Data is in contrast to that of intensely impulsive children who generate profiles marked by responses occurring between 1 or 2 seconds of the previous response. These are the more typically impulsive children who exhibit no resources for developing a strategy or for delaying at all.

Thus, the Task Monitoring Data can prove to be quite valuable in refining the overall interpretation. The information provided by the TMD not only helps in the diagnostic process, but can also provide critical input into the development of remedial and educational strategies.

INTERPRETATION OF THE VIGILANCE TASK

The Vigilance Task yields three kinds of Summary Scores: the number of Correct Responses (CR), the number of Errors of Omission (EO), and the number of Errors of Commission (EC).

Correct Responses (CR)

	6–7 YEAR OLDS	8–9 YEAR OLDS	10–11 YEAR OLDS
NORMAL	⩾23	⩾37	⩾41
BORDERLINE	7–22	28–36	38–40
ABNORMAL	0–6	0–27	0–37

This score reflects the total number of "1/9" combinations that were responded to by the subject. The sequence of digits presented is fixed and there are 15 "1/9" presentations in each 3-minute time block, for an overall total of 45 possible CR for the 9-minute session. The CR score reflects the subject's level of "vigilance" or alertness. In this respect, the term vigilance refers to the ability to focus the attentional processes in a goal-directed manner and to maintain this investment of attention over time. The CR score is an index of the subject's ability to achieve and maintain alert, vigilant responding. Children with CR scores in the normal range are thus able to maintain vigilance and attention appropriately, while children with BORDERLINE and ABNORMAL CR scores are those who tend to experience deficits in sustaining attention. It should be pointed out that a CR score within the ABNORMAL range, especially as unusually low as 4 or 5, may be due to the child's lack of motivation to maintain alertness. For this reason, we often suggest that these children be closely evaluated for emotional/motivational difficulties.

Interpretive Statements

Correct Responses within NORMAL range:

Subject is capable of investing sufficient energy in the attentional processes and maintaining this investment over time. S/he demonstrates intact vigilance under conditions of relatively high arousal. If this child was referred for inattentiveness, a more extensive psychological evaluation should be considered because problem behaviors do not appear to stem from a fundamental deficit in sustained attention. Anxiety, emotional difficulties, and/or environmental stress may account for these behaviors.

Correct Responses within BORDERLINE RANGE:

This subject's vigilance and alertness fall within the BORDERLINE range. Lapses in sustained attention are evident in his/her performance. Remedial strategies should involve helping him/her to remain attentive and to discriminate salient from non-salient stimuli. S/he may need instructions simplified and repeated. During task performance this youngster may require more supervision and monitoring than peers. Pharmacotherapy may be helpful as an augmentative intervention.

Correct Responses within ABNORMAL range:

This subject demonstrates significant deficits in sustained attention, which are possibly related to emotional or motivational difficulties. S/he is not alert to important cues, and may frequently miss or overlook salient aspects of a stimulus situation. This child might benefit from the "cognitive focusing" effect that has been noted with regimens of stimulant medication. Augmentative interventions should also include exercises geared toward learning attentiveness in high-arousal situations, such as flash-card exercises.

Errors of Omission (EO)

This score reflects the number of "1/9" presentations to which no response was made. Since the CR score reflects the number of these combinations to which the subject responded, the CR and EO scores mirror each other. For example, if a subject obtained a CR score of 41, then his or her EO score would have to be 4 (41 correct out of a possible 45). Consequently, the EO and CR scores are two sides of the same coin and receive identical interpretations. As with the CR score, the EO score is an index of lapses in this alert, vigilant attitude. These errors reflect instances in which the subject was not attending to the situation sufficiently to respond to the "1/9" sequence.

Errors of Commission (EC)

	6–7 YEAR OLDS	8–9 YEAR OLDS	10–11 YEAR OLDS
NORMAL	0–14	0–11	0–6
BORDERLINE	15–49	12–29	7–16
ABNORMAL	⩾50	⩾30	⩾17

The EC score reflects the number of times the subject responded to incorrect stimuli, that is, stimuli other than a

284

"9" immediately preceded by a "1". These errors are largely unrelated to either Correct Responses (CR) or Errors of Omission (EO), and actually constitute data on a separate dimension. While CR and EO reflect the adequacy of sustained attentional capabilities, EC reflect the degree of impulsivity the subject exhibits under the structure imposed by the cognitive demands of this task. The EC represents an index of the subject's inappropriate responding; i.e., responses which reflect not lapses in sustained attention, but rather responses to stimuli in an ill-conceived or poorly controlled fashion. While we will be discussing different categories of EC which reflect various aspects of a subject's functioning, keep in mind that all EC represent impulsive responding in one way or another.

Interpretive Statements

EC within NORMAL range:

Under conditions demanding vigilance and alertness, this subject demonstrates a level of impulsivity which is no different from his/her peers.

EC within BORDERLINE range:

Under conditions demanding vigilance and alertness, this subject demonstrates a level of impulsivity which is somewhat greater than his/her peers. S/he is likely to become disorganized during competitive classroom exercises, particularly when speed of responding is a requisite. Activities such as these should be avoided with this child until interventions geared toward reducing impulsivity under these circumstances can be effected.

EC within ABNORMAL range:

Under conditions demanding alertness, this subject demonstrates a high level of impulsivity. S/he is very likely to become disorganized during competitive classroom exercises, particularly when speed of responding is required. Competitive activities should be avoided. Evaluation for pharmacotherapy should be considered. As with many other very extreme GDS scores, these scores may reflect, in addition to a possible impulsive style, difficulties in motivation, lack of comprehension of the instructions, or interference by emotional factors. Further psychological evaluation is often recommended.

Errors of Commission fall into two basic categories which can be determined by inspection of the Task Monitoring Data:

Target-related Errors: These are the most frequent Errors of Commission. A Target-related Error is registered by any response to either the target stimulus ("9") or to the alerting stimulus ("1") when they are not in the correct "1/9" sequence. For example, a response to the alerting stimulus prior to the presentation of the next digit is a Target-related error. By the same token, a response to the target stimulus when it is not immediately preceded by the alerting stimulus also registers a Target-related error. A response to the "1" before the presentation of the next digit suggests that the subject is unable to inhibit responding long enough to await the next stimulus and scan it for correctness. A response to any non-target stimulus which immediately follows a "1" reflects that once the subject has been "cued" or primed to respond by the alerting stimulus he or she is unable to refrain from emitting a response, even if it is incorrect to do so. Similarly, responses to the "9", when not immediately preceded by the "1", reflect a disregard for the salient aspects of the stimulus situation.

Sometimes the subject responds correctly to the "1/9" sequence, but then responds to the very next digit presented, whatever that particular number may be. These errors reflect the inability to inhibit a response immediately after emitting a correct response. Most often, they relate to an elevated level of arousal which cannot be brought under control sufficiently to prevent an immediate, impulsive second response.

Random Errors: A Random Error is any button press that does not occur in response to either the "alerting stimulus" or to the "target stimulus". Thus, when the subject presses the button in response to a "4" or to a "7", when these stimuli were not immediately preceded by the "1" or "9", a Random Error is registered. These errors are highly unusual, and generally reflect a state of arousal which is out of control. However, a caveat applies here. In a protocol with many of these responses, the subject's motivation to perform well and comprehension of the instructions must be examined.

Block Scores

Just as with the Delay Task, the Block Scores reflect the overall variability in a subject's performance. This variability should be interpreted in the context of the total scores. For example, if the Total CR and EC are both NORMAL, the significant variability takes on a different meaning than when the Total Scores are BORDERLINE or ABNORMAL.

285

For the Vigilance Task Block Scores we examine two parameters, namely Slope and Extreme Scores (Peaks and Valleys):

Slope

As with the Delay Task we look for trends or patterns across the three blocks of the session. Most children maintain a consistent performance (Even Slope), but some improve (Upward Slope) and a few deteriorate (Downward Slope). These patterns can be estimated by visual inspection of the scores. For statistical purposes, we define a protocol as having an Upward Slope when each block has a greater number of responses (Correct or Commission) than the previous one. Conversely, a Downward Slope is when each block has fewer responses (Correct or Commission) than the preceding block. Keep in mind that an Upward Slope for CR reflects improvement while for EC it indicates deterioration.

Interpretive Statements

Even Slope for CR:
Subject demonstrates consistent performance across time.

Upward Slope for CR:
Subject initially demonstrates lapses in sustained attention, but performance improves over time as s/he becomes more familiar with the task.

Downward Slope for CR:
This subject demonstrates an initial ability to sustain attention and respond adequately, but his/her performance deteriorates over time.

Even Slope for EC:
This subject demonstrates consistent performance with respect to behavioral suppression.

Downward Slope for EC:
Performance improves over time with respect to behavioral suppression. In early stages of a task, performance may be disorganized and haphazard but becomes better controlled as subject adjusts to task demands.

Upward Slope for EC:
Performance deteriorates with respect to behavioral suppression. Subject becomes less controlled as task progresses.

Peaks and Valleys

In a protocol which is significant for deficits in sustained attention and/or behavioral suppression (CR and/or EC in ABNORMAL range), any Block Score along these two dimensions which is within normal limits must be considered in the interpretation.

Interpretive Statements

Total CR in BORDERLINE or ABNORMAL range and any Block CR greater than 8 for 6 & 7 year-olds and greater than 11 for the older groups:
While this subject demonstrates overall deficiencies in sustained attention, there are times when s/he is capable of sustaining attention as well as his/her peers. However, this ability is not of sufficient strength to bring overall functioning within the normal range.

Total CR in NORMAL range and any Block CR less than 4 for 6 & 7 year-olds and less than 9 for the older groups:
While this subject demonstrates essentially normal capacities for sustained attention, there are times when his/her attentional processes will falter to a level below that of his/her peers.

Total EC in BORDERLINE or ABNORMAL range and any Block EC less than 5:
Subject demonstrates capacity for normal behavioral suppression.

Total EC in NORMAL range and any Block EC greater than 10:
While subject demonstrates generally normal behavioral suppression, there are times when s/he will become impulsive when required to sustain attention.

RELATIONSHIP OF THE EC SCORE TO DELAY TASK SCORES

Interpretations of impulsivity based solely on Vigilance Task performance must be made with extreme caution. The GDS Vigilance Task places specific cognitive demands upon the subject, and the emergence of impulsive responding under these conditions may not necessarily reflect a core dimension of impulsivity. The Vigilance Task may stand alone when addressing issues related to the attentional processes and the ability to sustain attention. However, the interpretation of EC scores rests heavily upon the subject's performance on the Delay Task. The Delay Task taps into a central dimension of impulsivity, while EC scores reflect impulsivity which emerges or is suppressed under certain conditions.

Examples of Vigilance Task Interpretation

Three profiles are presented below to illustrate interpretation of the Vigilance Task. The Task Monitoring Data are displayed in a table that lists the frequency with which the subject responded to the various combinations of digits. In this table an "X" represents any digit that may have appeared. For example, the "X19" sequence indicates that the button was pressed upon presentation of a "9" which had been preceded by a "1", which in turn had been preceded by any number ("X"). This row would therefore show the number of Correct Responses. The "XX1" sequence records the number of times the button was pressed upon presentation of a "1" preceded by any two other numbers.

SUBJECT 5 – 9 year-old male

BLOCK	1	2	3	TOTAL
CORRECT	13	10	7	30
OMISSIONS	2	5	8	15
COMMISSIONS	14	23	37	74

SEQUENCE BLOCK:	1	2	3	TOTAL
X19	13	10	7	30
XX1	1	2	9	12
19X	1	2	3	6
XX9	0	0	0	0
X9X	2	6	4	12
X1X	7	4	6	17
XXX	3	9	15	27

This youngster evidences many of the hallmarks of an impulsive style. The Total CR of 30 is in the ABNORMAL range and indicates significant deficits in his ability to sustain attention in a goal-directed manner. He is apt to be quite inattentive in situations which require vigilance, such as a classroom setting. Block Scores indicate that, while his initial investment of attention is adequate, it deteriorates steadily as the task progresses. The EC score also suggests an extreme level of impulsive responding in situations requiring vigilance. As his capacity for sustained attention diminishes, his level of impulsive responding increases.

The Task Monitoring Data indicates that the majority of this child's Errors of Commission were "Target-related". This suggests that much of his impulsivity is related to his level of arousal: when he is "primed" to respond by the target stimulus he is unable to inhibit sufficiently to insure the appropriateness of his response. The Task Monitoring Data also reveal another important aspect of his performance. Note that in the first Time Block there were only 3 Random Errors (see the "XXX" row on the table); most of his errors were "Target-related". As the task progressed, his level of arbitrary responding (i.e., Random Errors) increased dramatically. This suggests that, as he becomes less able to sustain attention, he becomes more arbitrary and uncontrolled. It would be important to examine his Delay Task performance to determine his level of impulsivity in a self-paced task where arousal level becomes less critical.

287

SUBJECT 6 – 11 year-old boy

BLOCK	1	2	3	TOTAL
CORRECT	6	8	12	26
OMISSIONS	9	7	3	19
COMMISSIONS	16	8	5	29

SEQUENCE BLOCK:	1	2	3	TOTAL
X19	6	8	12	26
XX1	9	1	3	13
19X	6	0	1	7
XX9	0	0	0	0
X9X	0	0	0	0
X1X	1	7	1	9
XXX	0	0	0	0

This subject demonstrates deficits in sustained attention as evidenced by his ABNORMAL CR score of 26. He will therefore be apt to miss salient aspects of a situation. The Block Scores indicate that his inattentiveness and his impulsivity, both of which were problematical initially, improved over time. As such, this youngster will tend to be disorganized and inattentive upon the initial presentation of a task but can consolidate his performance as he "settles in".

A review of the Task Monitoring Data indicates that, while he is impulsive in the early stages of the task, his responding never became arbitrary (i.e., there were no Random Errors). The bulk of his errors were Target-related suggesting that once primed to respond, he is relatively unable to modulate his arousal in order to suppress inappropriate responses. He is thus likely to become over-stimulated when faced with competitive tasks or when speed is a major task requirement. Examination of his Delay Task performance would clarify the degree to which this impulsivity is manifested in a low-arousal situation.

SUBJECT 7 – 7 year-old girl

BLOCK	1	2	3	TOTAL
CORRECT	13	15	13	41
OMISSIONS	2	0	2	4
COMMISSIONS	2	1	0	3

SEQUENCE BLOCK:	1	2	3	TOTAL
X19	12	15	13	41
XX1	0	0	0	0
19X	0	0	0	0
XX9	0	0	0	0
X9X	0	1	0	1
X1X	2	0	0	2
XXX	0	0	0	0

This child demonstrates unusually well-developed capacities for sustained attention and behavioral suppression in situations requiring vigilance. Such a "clean" protocol is unusual for a youngster of her age, particularly since she was referred for an evaluation because of inattentiveness and poor self-control. Her Delay Task scores were also well within the NORMAL range. It was therefore suggested that the practitioner look toward motivational/emotional or environmental factors to account for the presenting complaints.

INTEGRATION OF VIGILANCE AND DELAY TASK SCORES

Although the two tasks contained within the GDS tap into different aspects of an individual's functioning, these dimensions are in many ways related. However, there will be many instances, particularly among clinical populations, when scores on the two tasks appear to contradict each other. The purpose of this section is to assist you in formulating interpretations when the two tasks are in accordance as well as when they are in apparent disagreement.

Toward this end, there are some points which must be addressed before interpretive statements are offered. Some of what follows has been covered in the GDS Interpretive Guide. However, a thorough understanding of these issues is essential to valid and effective interpretation.

The Delay Task is, first and foremost, a measure of behavioral suppression and impulse inhibition (delay). It places very few demands upon the subject's alertness or upon his/her ability to sustain attention. A very important aspect of the Delay Task is that it provides information or feedback to the subject. The flashing light and incrementing counter serve the purpose of providing the subject with a "feedback loop". For optimal performance, the subject must incorporate this feedback and use it to monitor and guide responding. In order to do this, most non-impulsive subjects develop cognitive "mediational strategies", usually involving some covert, verbally-mediated behaviors such as counting. Others develop behaviorally-mediated strategies, usually involving some regular pattern of motoric activity. We have observed that these children tend to make less efficient adaptations to this task than do those who employ more sophisticated, covert strategies.

It is important to note that these strategies are self-generated. The Delay Task is essentially subject-paced, in that the apparatus does nothing but record the subject's responding. The subject must pace his or her responding via the employment of these strategies. There is no aspect of the task which requires the subject to "gear up" to respond; therefore, any impulsivity which emerges does so in the absence of external pressure to respond in a rapid or hasty manner.

By contrast, the Vigilance Task requires a relatively high level of arousal on the part of the subject. The subject must achieve and maintain a stance of alertness and vigilance to perform well on this task. This demand is placed upon the subject by the apparatus.

As opposed to the self-paced nature of the Delay Task, the Vigilance Task is instrument-paced. The stimuli rapidly flashing on the screen require that the subject "gear up" for the possibility of responding with the presentation of each new stimulus. In addition, subject performance is in direct response to the instrument. There are no strategies needed to control and monitor responding; the subject simply must be prepared to respond to the correct stimuli at all times during the task.

This level of responding must be maintained in the absence of feedback. The Vigilance Task provides no information whatsoever on the appropriateness of the subject's responding. There is thus no ongoing process of cognitive mediation, as there is in the Delay Task. The subject must simply maintain a level of arousal or vigilance in order to maintain a state of "preparedness to respond". Obviously, optimal performance requires that the subject also suppress tendencies to respond impulsively while in this state of arousal.

Interpretive Statements

NORMAL Delay Task, NORMAL CORRECT (CR) & ERRORS OF COMMISSION (EC) on the Vigilance Task:

	Delay	Vigilance	
	ER	CR	EC
NORMAL	X	X	X
ABNORMAL			

This subject demonstrates an ability to delay and inhibit behavioral responding which is commensurate with his/her peers. This ability to delay is not only adequate in situations when behavioral suppression is the primary requirement, but also when high levels of arousal must be maintained. S/he can perform effectively when alertness and arousal are required, with no evidence of increased impulsivity under these conditions. If this child was referred for ADD, a more extensive psychological evaluation should be considered because problem behaviors do not appear to stem from fundamental deficits in impulse inhibition or sustained attention. Anxiety, emotional difficulties, and/or environmental stress may account for the presenting problems.

289

NORMAL Delay Task, NORMAL CR & ABNORMAL EC on Vigilance Task:

	Delay ER	Vigilance CR	Vigilance EC
NORMAL	X	X	
ABNORMAL			X

This subject demonstrates an ability to inhibit impulses which is commensurate with his/her peers in situations where behavioral suppression is required for successful performance. S/he can pace and monitor responding under conditions of low arousal and in the presence of informational feedback. The ability to sustain alertness and vigilance is also commensurate with peers. However, under conditions of high motivation & arousal where there is no informational feedback, responding becomes impulsive and poorly controlled. Programming should involve avoiding competitive, high-arousal activities in the classroom. Accuracy should be emphasized over speed of responding, and reinforcement should be used judiciously, as it may heighten arousal levels.

NORMAL Delay Task, ABNORMAL CR & EC on Vigilance Task:

	Delay ER	Vigilance CR	Vigilance EC
NORMAL	X		
ABNORMAL		X	X

This subject demonstrates an ability to inhibit or suppress behavioral responding in self-paced activities where feedback is provided. However, under conditions of high arousal and no feedback s/he demonstrates lapses in sustained attention or alertness in addition to abnormally high levels of impulsivity. In these situations s/he is prone to miss salient aspects of a situation and to respond in a poorly-controlled manner. S/he should be encouraged to utilize strengths (feedback utilization, self-pacing, delay), and interventions should be geared toward helping to incorporate these strengths in situations where vigilance is required. Competitive, speed-oriented exercises should be avoided. S/he should perhaps be allowed extra time to complete tasks such as exams and classroom quizzes.

NORMAL Delay Task, ABNORMAL CR & NORMAL EC on Vigilance Task:

	Delay ER	Vigilance CR	Vigilance EC
NORMAL	X		X
ABNORMAL		X	

This subject demonstrates normal capacity to delay and inhibit responding in situations where responding is self-paced and feedback is provided as well as in situations where high levels of arousal are required and no feedback is provided. However, s/he demonstrates lapses in alertness and sustained attention. S/he should be encouraged to utilize strengths (feedback utilization, self-pacing, delay), and interventions should be geared toward helping them to incorporate these strengths in situations where vigilance is required. S/he should perhaps be allowed extra time to complete tasks such as exams and classroom quizzes.

ABNORMAL Delay Task, NORMAL CR & EC on Vigilance Task:

	Delay ER	Vigilance CR	Vigilance EC
NORMAL		X	X
ABNORMAL	X		

This subject demonstrates a marked tendency to respond impulsively when s/he must pace and monitor responding. Feedback tends to be underutilized and mediational strategies tend to be ineffective. However, under conditions of increased structure (non-subject paced), s/he demonstrates adequate capacity for delay and normal attention and alertness. This child would likely benefit from self-control training programs aimed at teaching cognitively-oriented strategies for delay.

ABNORMAL Delay Task, NORMAL CR & ABNORMAL EC on Vigilance Task:

	Delay ER	Vigilance CR	EC
NORMAL		X	
ABNORMAL	X		X

This subject demonstrates significant impulsivity, which emerges across many situations. Although vigilance and alertness is commensurate with that of peers, impulsivity is pervasive and may profoundly affect efficiency and effectiveness. S/he will require greatly increased external structure in the classroom, particularly when confronted with tasks which involve self-control and the capacity to inhibit responding. This child would benefit from special programs designed to teach strategies for developing cognitive self-control over behavior. This child will require help in learning how to "stop, look, listen, and think" before acting. Treatment with stimulant medication may be a viable option to consider as part of a general program of educational and psychological management.

ABNORMAL Delay Task, ABNORMAL CR & NORMAL EC on Vigilance Task:

	Delay ER	Vigilance CR	EC
NORMAL			X
ABNORMAL	X	X	

This subject demonstrates a degree of impulsivity considerably greater than that of his/her peers. This tendency toward impulsive responding is most prominent in situations in which behavioral suppression is required and s/he must pace and monitor his/her responding. This tendency diminishes under conditions of increased structure (when tasks are not self-paced). However, under these conditions s/he exhibits marked losses in alertness and sustained attention. S/he will thus be prone to miss salient aspects of stimulus situations. Motivation to perform under conditions of high arousal is suspect, as subject may be inhibiting unduly to avoid a torrent of impulsive responding.

ABNORMAL Delay Task, ABNORMAL CR & EC on Vigilance Task:

	Delay ER	Vigilance CR	EC
NORMAL			
ABNORMAL	X	X	X

This subject demonstrates pervasive impulsivity coupled with deficits in sustained attention. S/he responds in an abnormally impulsive manner across situations, and is prone to miss the salient aspects of a stimulus situation. This child should be considered for treatment with stimulant medication. S/he will require greatly increased external structure in the classroom, particularly when confronted with tasks which involve self-control and the capacity to inhibit responding. Programming should avoid high-arousal situations and include training in cognitive self-control strategies.

The interpretations offered above can be embellished and refined considerably by taking into account the variety o factors within each task as well as by incorporating findings from other sources. Moreover, once the practitioner becomes thoroughly familiar with the nuances of GDS performance and scores, interpretations will take on a much more individualized character.

291

STANDARD INSTRUCTIONS

Delay Task

You're going to play a game in which you will get a chance to win a lot of points. Do you see this light (pointing to the small red light)? Every time you make this light go on you'll earn a point and this counter (pointing) will keep track of how many points you've won. At the end of the game we'll see how many points you've earned. Now, to make the light go on all you have to do is push this blue button (pointing), and wait a little while before pressing it again. You just press this blue button, wait a while, then press it again. If you press it again too soon, though, you will not get a point, the light won't go on, and you'll have to wait a while before you can press it to get another point But if you push the button, wait a while, then push it again, you'll get a point every time.

Vigilance Task

Following the Delay Task

Now you're going to play another game. In this game you will see numbers flash on the screen, and I want you to press the blue button every time you see a "9" that comes right after a "1". If the "9" comes after any other number, don't press the button. The only time you should press it is if you see a "9" that comes right after a "1". Now the red light won't go on at all but at the end of the game I'll tell you how many points you won. You will know the game is over when this green light goes on. Do you understand? (Have subject repeat.).

Not Following the Delay Task

This is a game in which you will see numbers flash on the screen quickly. What you need to do in order to win points is to press the blue button every time you see a "9" that comes right after a "1". If the "9" comes after any other number, don't press the button. The only time you should press it is if you see a "9" that comes right after a "1". You will know the game is over when this green light goes on. Do you understand? (Have subject repeat.)

GORDON DIAGNOSTIC SYSTEM

THRESHOLD TABLES

ALL AGES	NORMAL	BORDERLINE	ABNORMAL
Efficiency Ratio Slope	0 - ±.15		< -.15 or >.15

3-5 YEAR OLD CHILDREN	NORMAL	BORDERLINE	ABNORMAL
DELAY TASK (N=150)			
Efficiency Ratio			
Total	.61 - 1.00	.43 - .60	0 - .42
Block Variab.	.00 - 0.14	.15 - .25	≥ .26
Single Block	.61 - 1.00		0 - .40
Responses - Total	21 - 66	67 - 89	≤20 or ≥90
Correct - Total	≥ 26	16 - 25	0 - 15
VIGILANCE TASK (N=132)			
Correct			
Total	21 - 29	14 - 20	0 - 13
Single Block	NA		NA
Commissions			
Total	0 - 6	7 - 25	≥ 26
Single Block	0 - 1		≥ 11
Block Variab.	0 - 2.07	2.08-4.99	≥ 5

6-7 YEAR OLD CHILDREN	NORMAL	BORDERLINE	ABNORMAL
DELAY TASK (N=142)			
Efficiency Ratio			
Total	.72 - 1.00	.42 - .71	0 - .41
Block Variab.	.00 - 0.11	.12 - .20	≥ .21
Single Block	.71 - 1.00		0 - .39
Responses - Total	21 - 63	64 - 75	≤20 or ≥76
Correct - Total	≥ 35	15 - 34	0 - 14
VIGILANCE TASK (N=100)			
Correct			
Total	33 - 45	22 - 32	0 - 21
Single Block	12 - 15		0 - 6
Commissions			
Total	0 - 10	11 - 23	≥ 24
Single Block	0 - 3		≥ 9
Block Variab.	0 - 2.07	2.08-4.99	≥ 5

8-9 YEAR OLD CHILDREN	NORMAL	BORDERLINE	ABNORMAL
DELAY TASK (N=137)			
Efficiency Ratio			
Total	.78 - 1.00	.52 - .77	0 - .51
Block Variab.	.00 - 0.10	.11 - .23	≥ .24
Single Block	.77 - 1.00		0 - .49
Responses - Total	21 - 65	66 - 75	≤20 or ≥76
Correct - Total	≥ 41	22 - 40	0 - 21
VIGILANCE TASK (N=96)			
Correct			
Total	40 - 45	33 - 39	0 - 32
Single Block	14 - 15		0 - 11
Commissions			
Total	0 - 8	9 - 21	≥ 22
Single Block	0 - 2		≥ 9
Block Variab.	0 - 1.52	1.53-3.78	≥ 3.79

10-11 YEAR OLD CHILDREN	NORMAL	BORDERLINE	ABNORMAL
DELAY TASK (N=131)			
Efficiency Ratio			
Total	.81 - 1.00	.55 - .80	0 - .54
Block Variab.	.00 - 0.11	.12 - .25	≥ .26
Single Block	.80 - 1.00		0 - .47
Responses - Total	21 - 67	68 - 76	≤20 or ≥77
Correct - Total	≥ 45	32 - 44	0 - 31
VIGILANCE TASK (N=88)			
Correct			
Total	42 - 45	39 - 41	0 - 38
Single Block	14 - 15		0 - 12
Commissions			
Total	0 - 3	4 - 15	≥ 16
Single Block	0 - 1		≥ 5
Block Variab.	0 - 1.15	1.16-1.99	≥ 2

12-16 YEAR OLD CHILDREN	NORMAL	BORDERLINE	ABNORMAL
DELAY TASK (N=156)			
Efficiency Ratio			
Total	.86 - 1.00	.71 - .85	0 - .70
Block Variab.	.00 - 0.10	.11 - .20	≥ .21
Single Block	.86 - 1.00		0 - .66
Responses - Total	21 - 66	67 - 75	≤20 or ≥76
Correct - Total	≥ 47	39 - 46	0 - 38
VIGILANCE TASK (N=156)			
Correct			
Total	43 - 45	40 - 42	0 - 39
Single Block	15		0 - 12
Commissions			
Total	0 - 3	4 - 11	≥ 12
Single Block	0		≥ 5
Block Variab.	0 - .99	1.00-1.[illegible]	≥ 2

PRESCHOOL INSTRUCTIONS DELAY TASK

"You are going to play a game in which you will get a chance to win a lot of points, not just one or two, but a whole bunch of points. Do you see this light (pointing to the small red light)? Every time you make this light go on, you'll earn a point, and this counter (pointing) will keep track of how many points you've won. At the end of the game, we'll see how many points you've earned. Now, to make the light go on, all you have to do is push this blue button (pointing), and wait a little while, then press it again. If you press it again too soon, though, you will not get a point, the light won't go on, and you'll have to wait a while before you can press it to get another point. But, if you push the button, wait a while, then push it again, you'll get a point every time.

Do you have any questions?

I'll sit back here and wait until you've finished. You will know the game is over when the green light comes on and a buzzer goes off. If you have any questions or want to talk about the game, I want you to wait until the game is over, and we can discuss it then.

Are you ready (said with enthusiasm)? OK, let's begin (or "ready, set, go!").".

VIGILANCE TASK

Instruction And Practice Session

"Now to play the next game, you need to know what the number "1" looks like. Can you show me a "1" (showing them the flash card with LED-style numbers)? Can you show me a "7"? Can you show me a "1" again? (The child is asked to point to the number "7" because it looks similar to a "1", so that it is necessary to make sure that the child can differentiate between the two numbers. Since some children are unable to discriminate between the numbers, you should try to teach them. If the child still cannot pick out the "1", then the task cannot be administered.)

When you play this game, you are going to see numbers flash on the screen, and I want you to press the blue button only when you see a "1". Only press the blue button if you see a "1" and not when any other number flashes. Are you ready? Begin."

Assessment Session

"Now you are going to play this game for a longer time. I want you to press the blue button only when you see a "1". Only press the button if you see a "1". I'll sit here and wait until you're done. You'll know when the game is over, because the green light will come on, and you will hear a beep. If you have any questions, or want to talk about the game, I want you to wait until after the game is over, and we can discuss it then."

DELAY TASK SETTINGS

- *Set TASK SELECTOR SWITCH to **DELAY TASK***
- ***INTERVAL** Thumbwheel - 4 seconds*
- ***BLOCK LENGTH** Thumbwheel - 90 seconds*
- *Press four Block **PROGRAMMER** buttons*

ONE MODE/VIGILANCE TASK SETTINGS

Instruction And Practice Session

- *Set TASK SELECTOR SWITCH to **VIGILANCE TASK***
- *Set **Mode Selector** to "1"*
- ***INTERVAL** Thumbwheel - 2 seconds*
- ***BLOCK LENGTH** - 20 seconds*
- *Press **BLOCK 1** Programmer button*

  a. ***BLOCK LENGTH** Thumbwheel - 0 seconds*
  b. ***INTERVAL** Thumbwheel - 2 seconds*
  c. *Press **BLOCK 2** and **3** Block Programmer buttons*

Assessment Session

- *Press **RESET** button*
- ***BLOCK LENGTH** Thumbwheel - 120 seconds*
- ***INTERVAL** Thumbwheel - 2 seconds*
- *Press first 3 **BLOCK** Programmer buttons*

REFER TO MANUAL FOR DETAILED INSTRUCTIONS

GORDON DIAGNOSTIC SYSTEM
PRESCHOOL CHILDREN
THRESHOLD TABLES

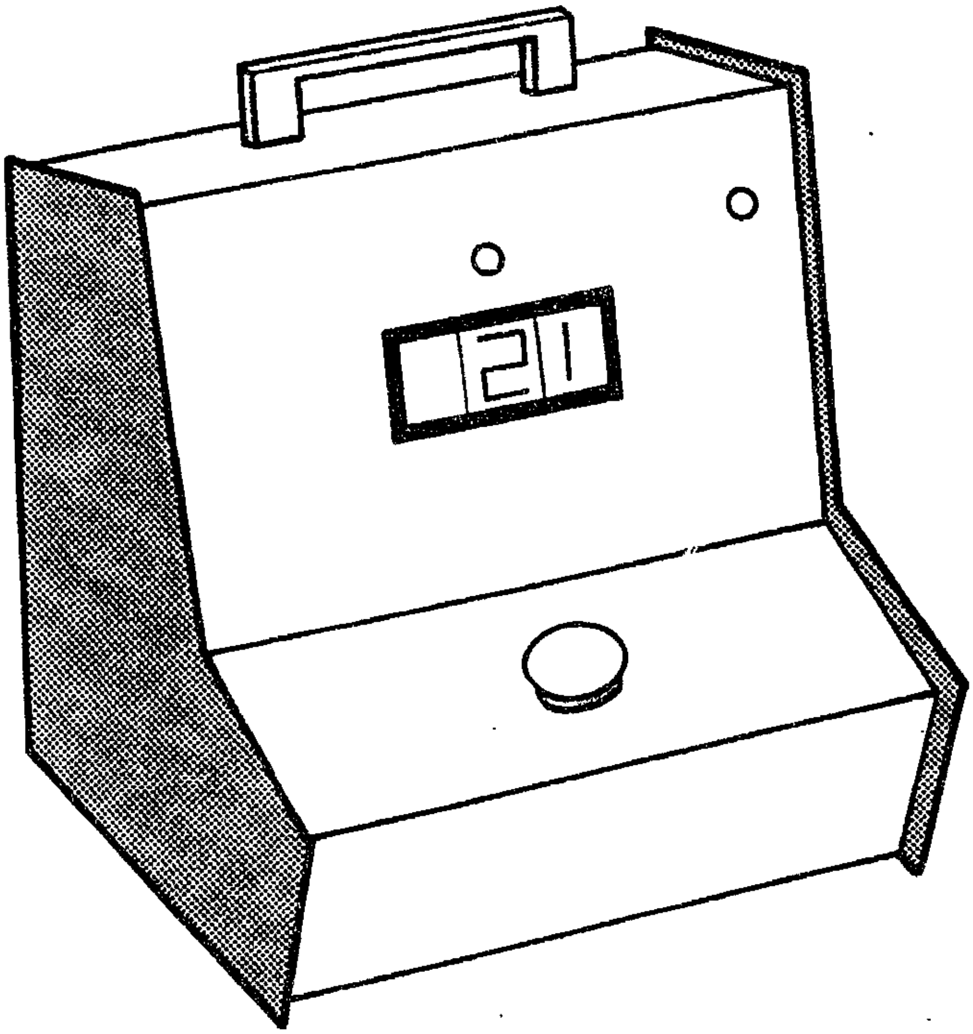
	NORMAL	BORDERLINE	ABNORMAL
Efficiency Ratio Slope	0 - ±.15		< -.15 or >.15

3-5 YEAR OLD CHILDREN	NORMAL	BORDERLINE	ABNORMAL
DELAY TASK (N=150)			
Efficiency Ratio			
Total	.61 - 1.00	.43 - .60	0 - .42
Block Variab.	.00 - 0.14	.15 - .25	≥ .26
Single Block	.61 - 1.00		0 - .40
Responses - Total	21 - 66	67 - 89	≤20 or ≥90
Correct - Total	≥ 26	16 - 25	0 - 15
VIGILANCE TASK (N=132)			
Correct			
Total	21 - 29	14 - 20	0 - 13
Single Block	NA		NA
Commissions			
Total	0 - 6	7 - 25	≥ 26
Single Block	0 - 1		≥ 11
Block Variab.	0 - 2.07	2.08-4.99	≥ 5

295

THE GORDON DIAGNOSTIC SYSTEM

DATA ANALYSIS PROGRAM



CLINICAL DIAGNOSTICS INC
1536 Cole Blvd., Golden, Colorado 80401
Suite 350 • 800-521-4503

THE GORDON DIAGNOSTIC SYSTEM DATA ANALYSIS PROGRAM

Ernest M. Post, M.D.
Cooper Hospital/University Medical Center, Camden, New Jersey

Michael Gordon, Ph.D.
S.U.N.Y. Upstate Medical Center, Syracuse, New York

Table of Contents

297

Organization of this Manual

The GDS Data Analysis Program (DAP) was developed to enhance and facilitate the operation of the Gordon Diagnostic System (GDS). This manual explains how to run the DAP on your computer. Since many DAP users have never before touched a computer, we have tried to keep the terminology and instructions as simple as possible.

It is not our intention to teach you to use your personal computer; that is the function of the operating manual provided with your machine. Most of the operation of the DAP is the same regardless of which computer you are using. Those features which vary among machines are mentioned briefly in the text and explained further in the Appendix.

I. **GETTING STARTED** tells you how to connect the GDS to your computer and how to start running the DAP.

II. **CONVERSING WITH THE DAP** describes in general terms how the computer asks for information and how you should respond.

III. **STEP BY STEP** tells you how to direct the DAP to collect, interpret, and save the results generated by the GDS.

IV. **FUNCTIONS MENU** presents the first half of the DAP features you can access using the function keys.

V. **STORED RESULTS** describes the other features of the DAP, which let you recall information stored during previous testing sessions.

APPENDICES

A. **SYSTEM REQUIREMENTS** is a general description of what you need to connect the GDS to a computer.

B. **SPECIFIC COMPUTERS** contains the necessary details for running the DAP on a variety of computers.

C. **ORGANIZATION OF DATA FILES** describes the file structure so that you may access the files from other programs.

D. **MAKING COPIES** is "must reading" for the computer novice. It discusses the need to keep duplicate "copies" of all important documents.

Acknowledgements

We would like to thank everyone who helped refine The Data Analysis Program and clarify this manual. We are particularly grateful for the contributions of Dr. F. Daniel McClure and the thorough evaluations by Ms. Barbara Bass and Ms. Mary Gallaghan. As always, we appreciate the gracious tolerance of our wives and children.

January 26, 1984

I. GETTING STARTED

Before you can start using the DAP, you need to have the appropriate pieces of equipment properly connected. The details are described in Appendix A: **SYSTEM REQUIREMENTS**, but basically you need the Gordon Diagnostic System (GDS), an appropriate computer, and the DAP. Depending on your computer, the DAP may have been provided for you on a "floppy disk", tape cassette, or "magnetic bubble" cartridge.

1. Setting up the Equipment

First, if your computer has separate components, you should connect them following the instructions in your owner's manual. Next, the GDS and the computer need to be connected by a cable (see Appendix A), which should have been purchased with your computer. (If the connectors on the two ends of the cable are the same, it doesn't matter which end goes where.) There is only one possible place to connect the cable to the GDS. Some computers may have more than one place where the other end of the cable could be connected; be sure you connect the cable to the part of the computer called the "serial port", the "asynchronous port", or the "RS-232 connector". (The three terms are essentially synonymous.)

2. Getting the DAP into your computer

Once you have everything connected, you need to "load" the DAP into your computer. If it is stored on a floppy disk, you need to insert the disk in your computer as described in the owner's manual. If the DAP is in a magnetic bubble cartridge, the cartridge needs to be in place in the computer. Unlike the floppy disk, it can be left in place when the machine is turned off, so it need not be inserted for each session. If the DAP is stored on a tape cartridge, it needs to be "loaded" from the cartridge to the computer before it can run. For some computers, once loaded, the DAP will stay in the computer and need not be loaded for each session.

3. Starting the DAP

As with the previous section, there are differences among computers as to the procedures for starting the DAP. In some computers, the DAP is set up to start running automatically. In others, you will need to tell the computer that you want to run the DAP, as opposed to some other program you may have. Specific details of how to start the DAP in each machine have again been left for Appendix B.

299

II. CONVERSING WITH THE DAP

The DAP lets you decide what to do by displaying a series of "prompts" or questions, to which you must respond by pressing one or more keys on the computer keyboard. The DAP only recognizes capital letters; if your computer has a lower case mode, you should press the "CAPS", "CAPS LOCK", or other appropriate key to make all the letters you type appear in UPPER CASE. (If you're unsure, see Appendix B.)

There are several types of messages (prompts) displayed by the DAP to request a response from you. Each is described below and followed by an example:

1. Yes/No Choices

These are generally self-explanatory. For example, the very first prompt to appear is:

PRINTER READY (Y/N)?>

You are expected to respond by pressing either the "Y" or "N" key on the computer keyboard. For this and several other responses only a single key needs to be pressed. To repeat, RESPONSES MUST BE IN CAPITAL LETTERS OR THE DAP WON'T ACKNOWLEDGE THEM.

2. Other two-option choices

These prompts are similar to the (Y/N) prompts. One prompt using this format is:

QUICK OR STANDARD LIST (Q/S)?>

You would respond with the appropriate letter, Q or S, as will be explained later.

3. Function Keys

If your computer (see Appendix B) has keys labelled F1, F2, F3, etc., these are called function keys because they are used to select the function which you wish the program to perform. The functions of the DAP will be explained in detail in the next section, but are listed below:

Key Function
F1 IDENTIFY NEW SUBJECT
F1 ENTER RESULTS
F3 REDISPLAY CURRENT RESULTS
F4 LIST STORED SUBJECTS
F5 RECALL STORED RESULTS
F6 DELETE STORED RESULTS
F7 END SESSION

There is one other function which doesn't appear on this list, but which can be implemented whenever this list is displayed by the computer (except on the Apple). As will be stated by the computer, you can begin running one of the GDS tasks; when the computer displays the list of functions it is also ready to start receiving information about the tasks from the GDS.

4. Print Option

The computer displays the results of the GDS testing and their interpretations on its screen in several segments as described in the next chapter. At the end of each of these segments the following prompt appears:

PRINT/CONTINUE/SKIP RESULTS (P/C/S)?

If you respond by pressing "P", and you previously told the DAP you had a printer ready, the computer will print exactly what just appeared on its screen to give you a paper copy for future reference. If you press "C", the computer will go on to the next step, which is often displaying more results. If you press "S", the computer will bypass displaying the remainder of the results and give you a choice of the "functions" again. In some instances, the pressing either "C" or "S" will have the same effect. You never need to worry that pressing "S" will result in losing data; it skips the display of the results, but does not affect their storage.

5. Requests for specific information

Sometimes the computer needs more than a single letter response from you. For example, when the computer wants to know the name of the examiner, the following prompt appears:

EXAMINER NAME>

You then type the examiner's name exactly as you want it to appear on everything that the computer will print about that particular subject. You may use last name first, abbreviations, or nicknames if you so choose. Whatever you enter is exactly what the computer will give back to you. For the computer to know that you have finished typing the name, YOU MUST PRESS the < ENTER > or < RETURN > key (sometimes marked by an arrow pointing down, then left). If you make a mistake while typing your response and you haven't yet hit the <ENTER> key, typing errors can be corrected by pressing the "Backspace" button. Its position and operation vary among computers; consult your computer manual to learn how to use it. Even if you have pressed <ENTER>, you will still have an opportunity to make changes later on.

6. Subject Name

There are several prompts similar to the one just described, but one is slightly different and is described below:

Request for subject name – This prompt can appear in two forms. When it appears as

SUBJECT NAME>

it is the same as other requests for information. What you get back will be exactly what you put into the computer. When it appears as:

ENTER SUBJECT NAME>

it is more flexible. This form is used when you are asking the computer to do something with data that has already been stored. You do not need to enter the subject's full name; any part of it will do. For example, if the subject is MIKE SMITH, you only need to answer this prompt with MIKE. The computer will search for the record of anyone named MIKE and ask you if it found the right Mike before it goes on. However, the computer cannot find the record if you originally entered the name as MICHAEL SMITH. If you forget how you entered it, you could ask for MI, but be careful because the computer will also find MIRIAM and TAMMI that way. If you ask for SMITH, it will try all subjects named SMITH until it finds the one you want.

7. Escape Key

Because it takes the computer several seconds to generate some of the displays described below, the "long" ones are provided with an "escape" mechanism. If you press the <ESCAPE> key (See Appendix B) the routine will be interrupted and the computer will go on to the next step. If you "escape" before or during the screen display, you will still have an opportunity to print the display on your printer. The displays which use this feature are the graphics, the Interpretations, and the Task Monitoring Data.

301

III. STEP BY STEP

This chapter gives detailed instructions for using the DAP. You may want to try each step as you read about it. The sections of this chapter are as follows:

1. Printer Ready Prompt
2. Starting a Task
3. Subject Identification
4. Data Display

1. Printer Ready Prompt

Once you have started the DAP, as described in GETTING STARTED, you should hear a beep and see:

GDS Data Analysis Program

(C) 1984 Clinical Diagnostics, Inc.
Serial #840120

USE CAPITAL LETTERS ONLY
PRINTER READY (Y/N)?>

Answer this question carefully. If the printer is connected and ready, type Y. If you type Y and the printer is not ready, the program may not be able to proceed. The following message will appear:

IF THIS STAYS ON PRINTER'S NOT READY.

PRESS <BREAK> THEN <F8>

Pressing <BREAK> can require that you press two keys rather than one for some computers **(See Appendix B).**

If you type N in response to the PRINTER READY prompt, the DAP will not print anything on the printer until you have started again from the beginning. If you don't have a printer ready when you are administering the GDS, don't worry. The data collected from the GDS can be saved and printed later if desired.

2. Starting a Task

Once you have responded to the PRINTER READY prompt, the following will appear on your computer:

START TASK OR PRESS A FUNCTION KEY
F1 IDENTIFY NEW SUBJECT
F2 ENTER RESULTS
F3 REDISPLAY CURRENT RESULTS
F4 LIST STORED SUBJECTS
F5 RECALL STORED RESULTS
F6 DELETE STORED RESULTS
F7 END SESSION

The computer is now ready to accept data from the GDS. You do not have to do anything else with the computer itself. Simply begin the GDS task, according to procedures outlined in the GDS Instruction Manual. (With the Apple computer, you will need to press F2 and then E before the computer is ready to accept data from the GDS.)

While the task is being performed, Task Monitoring Data will appear on the computer screen. You may observe the progress of the task, but it may be more valuable to keep your attention on the child's behavior at this time.

The data that is displayed is different for the two tasks. For the Delay Task the following is displayed each time the blue button is pressed:

INTERVAL #n WAS x.x SEC

This tells you how long it was since the previous button press and gives a picture of the child's progress. For the Vigilance Task, the display is a series of digits separated by a blank space or by a hyphen (dash). The digits are those displayed on the GDS for the task. The presence of a dash indicates that the blue button was pressed after that digit. If there is no dash, the button was not pressed.

While observing these results can be informative, there is no need to take notes because a summary of them is displayed by the computer at the conclusion of the task. (See Data Display below.)

302

3. Subject Identification

When the task has been completed, i.e., when the GDS beeps and the green Game Over light comes on, the following prompt will appear:

nn OF mm POSSIBLE SUBJECTS STORED

SUBJECT NAME>

The first line tells you how much storage space you have left. If your space is full, you will be so informed and offered the opportunity to delete a subject **(See Chapter V. F6 DELETE STORED RESULTS)**.

In response to the SUBJECT NAME prompt you should enter the name of the subject exactly as you want it to appear when the computer displays or prints the task results. You may use last name first, with or without a comma. You are limited to a maximum of 25 letters, spaces, numbers or punctuation marks (collectively termed characters). After you type the name, press the <ENTER> key to tell the computer you have finished entering the name. The computer will then ask:

SEX (M/F)>

You should type M or F, and then press <ENTER>. The next prompt is:

BIRTH (MM/DD/YY)>

to which the appropriate response would be 05/09/76 for a child born on May 9, 1976. The computer does not check that you have entered the date properly, but the zeros before the 5 and 9 are needed for the computer to determine the child's age. Next is:

EXAMINER NAME>

and again you enter the name as you wish, with a maximum of 13 characters. After that comes:

COMMENT>

Here you can respond with up to 13 characters of information (e.g., you may want to identify the child's school or reason for referral). The last prompt in this series is:

TEST DATE (MM/DD/YY)>

You need to respond in the same format as that used for the birth date. The computer then asks:

CHANGE THESE ENTRIES (Y/N)?>

If you answer Y, then all of the subject identification prompts are repeated and all the information needs to be reentered. If you answer N, your responses are stored and the computer goes on to the next step.

You may leave the sex, examiner and comment entries blank by pressing <ENTER> without typing anything. If you leave either date blank, the computer will be unable to determine the subject's age. If you leave the subject's name blank, the data CANNOT be stored. The results will still be displayed, but cannot be retrieved later on.

There is an earlier opportunity to enter all of the subject identification data, but that won't be discussed until **Chapter IV, Functions Menu.**

4. Data Display

After you have administered the task and entered the subject information, the computer automatically starts to display the task results and their interpretation. Samples of each display are shown and described below. The first display repeats the subject identification information to give you an opportunity to print it.

```
                    SUBJECT NUMBER 6

SUBJECT NAME:                          MICHAEL SMITH
SEX (M/F):                                         M
BIRTH (MM/DD/YY):                           05/09/76
EXAMINER NAME:                           JOHN PSYCHE
COMMENT:                                   R/O A.D.D.
TEST DATE (MM/DD/YY):                       01/10/84
PRINT/CONTINUE/SKIP RESULTS (P/C/S)?
```

Choose one of the alternatives, P. C, or S, as described in **Chapter II, Conversing with the DAP.**

The next display shows the Summary Statistics, which includes the Parameters used when administering the task and the scores achieved on the task. These scores are the same as those that can be read from the back panel of the GDS using the Data Selector. The displays from the Delay and Vigilance Tasks are slightly different, so only the Delay Task is illustrated here. An illustration of the Vigilance Task display is included at the end of this chapter.

BLOCK:	1	2	3	4	TOTAL
INTERVAL	6	6	6	6	
LENGTH	120	120	120	120	480
REWARDS	10	8	10	4	32
RESPONSES	18	23	23	27	91
EFFIC RATIO	.56	.35	.43	.15	.35

All of the terms in the above example are described in the Instruction Manual, Interpretive Guide, and Interpretive Supplement. To briefly review, the Interval and Length are the Task Parameters you specified. Rewards is the number of button presses that were made after waiting a sufficient period of time. Responses is the total number of button presses and Efficiency Ratio is determined from dividing the number of Rewards by the number of Responses. These results are displayed for each time Block and for the total Task.

The next display is a graphic presentation of the Summary Statistics.

```
                         REWARDS
      !                                             !
 25 _ !                                             !
      !                                             !
 20 _ !                                             !
      !                                             !
 15 _ !                                             !
      !                                             !
 10 _ !                                             !
      !   XXX       XXX       XXX                   !
  5 _ !   XXX       XXX       XXX                   !
      !   XXX       XXX       XXX       XXX         !
  0 _ !   XXX       XXX       XXX       XXX         !
       ------------------------------------------------
           1         2         3         4
                         BLOCKS

                        RESPONSES
      !                                             !
 50 _ !                                             !
      !                                             !
 40 _ !                                             !
      !                                             !
 30 _ !                                             !
      !                                 XXX         !
 20 _ !             XXX       XXX       XXX         !
      !   XXX       XXX       XXX       XXX         !
 10 _ !   XXX       XXX       XXX       XXX         !
      !   XXX       XXX       XXX       XXX         !
  0 _ !   XXX       XXX       XXX       XXX         !
       ------------------------------------------------
           1         2         3         4
                         BLOCKS

                       EFFIC RATIO
      !                                             !
  1 _ !                                             !
      !                                             !
 .8 _ !                                             !
      !                                             !
 .6 _ !                                             !
      !   XXX                                       !
 .4 _ !   XXX                 XXX                   !
      !   XXX       XXX       XXX                   !
 .2 _ !   XXX       XXX       XXX                   !
      !   XXX       XXX       XXX       XXX         !
  0 _ !   XXX       XXX       XXX       XXX         !
       ------------------------------------------------
           1         2         3         4
                         BLOCKS
```

For each graph, the bar extends above the horizontal line at the left if the value exceeds the corresponding number. For example, the Rewards for Block 1 was 10, which is greater than 7.5, but not greater than 10, so that the bar extends up to, but not above the line at 10. The graphics can be bypassed by pressing the <ESCAPE> key.

The next display shows interpretations based upon the subject's scores. (For the Vigilance Task, the interpretations also depend on the subject's age.) The logical basis of the interpretation is described in the Supplement to the Interpretive Guide.

TOTAL SCORES INDICATE:
Significant inability to delay
Shows clear evidence of impulsivity
Interpret protocol with extra caution because of very high responsivity

BLOCK SCORES INDICATE:
Never reaches normal range of behavioral delay
Block variability in borderline range
May be more inconsistent than peers in level of impulsivity exhibited
Performance deteriorates during session
Will become more impulsive over time

EDUCATIONAL PROGRAMMING:
Teach cognitive mediational strategies
Emphasize correctness over speed
Encourage self-monitoring
Evaluate for possible motivational/emotional factors
Extremely unable to suppress responding
Will need increasing structure as task progresses

The interpretations are only available if the Standard Parameters were used and, for the Vigilance "1/9" task if the child is at least 6 years old. This routine can be bypassed by pressing the <ESCAPE> key.

The final display is a summary of the Task Monitoring Data and gives the details of the subject's performance on the task. For the Delay Task it appears as follows:

INITIAL RESPONSE LATENCY = 01.7 SEC

INT.	BLOCK: 1	2	3	4	TOTAL
0 SEC	0	0	0	0	0
1 SEC	1	0	0	0	1
2 SEC	1	2	1	2	6
3 SEC	4	4	3	6	17
4 SEC	1	5	5	9	20
5 SEC	1	4	4	6	15
6 SEC	4	3	4	2	13
7 SEC	4	3	5	0	12
8 SEC	1	1	1	1	4
9 SEC	0	1	0	0	1
10 SEC	0	0	0	0	0
11 SEC	0	0	0	0	0
12 SEC	0	0	0	1	1
13 SEC	0	0	0	0	0
14 SEC	0	0	0	0	0
15 SEC	0	0	0	0	0
16 SEC	0	0	0	0	0
17 SEC	0	0	0	0	0
18 SEC	0	0	0	0	0
19 SEC	0	0	0	0	0
>19 SEC	0	0	0	0	0

This display tells you the number of times the button was pressed after the stated interval, for each Block and for the total Task. This can be bypassed by pressing the <ESCAPE> key.

After this display, the computer returns to the same display (the Functions Menu) as that shown in **Section 2, Starting a Task** and which is explained further in the next chapter.

305

Vigilance Task Display Example

SUMMARY STATISTICS

BLOCK:	1	2	3	TOTAL
INTERVAL	1	1	1	
LENGTH	180	180	180	540
CORRECT	12	7	11	30
OMISSIONS	3	8	4	15
COMMISSIONS	7	18	21	46

The terms in the above example that differ from the Delay Task are: Correct, Omissions, and Commissions. As explained in the Interpretive Guide, Correct is the number of times the button was pressed after the 1/9 sequence appeared. Omissions is the number of times the subject erred by missing the 1/9 sequence. Commissions is the number of errors the subject made by pressing the button when the 1/9 sequence had not appeared.

GRAPHICS

The displays for the Vigilance Task are similar to those generated for the Display Task.

```
                        CORRECT
      !                                          !
 15 _ !                                          !
      !                                          !
 12 _ !                                          !
      !   XXX                         XXX        !
  9 _ !   XXX                         XXX        !
      !   XXX                         XXX        !
  6 _ !   XXX           XXX           XXX        !
      !   XXX           XXX           XXX        !
  3 _ !   XXX           XXX           XXX        !
      !   XXX           XXX           XXX        !
  0 _ !   XXX           XXX           XXX        !
      --------------------------------------------
           1             2             3
                       BLOCKS

                       OMISSIONS
      !                                          !
 15 _ !                                          !
      !                                          !
 12 _ !                                          !
      !                                          !
  9 _ !                                          !
      !                 XXX                      !
  6 _ !                 XXX                      !
      !                 XXX                      !
  3 _ !                 XXX           XXX        !
      !   XXX           XXX           XXX        !
  0 _ !   XXX           XXX           XXX        !
      --------------------------------------------
           1             2             3
                       BLOCKS

                      COMMISSIONS
      !                                          !
 30 _ !                                          !
      !                                          !
 24 _ !                                          !
      !                                          !
 18 _ !                               XXX        !
      !                 XXX           XXX        !
 12 _ !                 XXX           XXX        !
      !                 XXX           XXX        !
  6 _ !   XXX           XXX           XXX        !
      !   XXX           XXX           XXX        !
  0 _ !   XXX           XXX           XXX        !
      --------------------------------------------
           1             2             3
                       BLOCKS
```


306

INTERPRETATIONS

SUBJECT'S AGE 7.5 YEARS

TOTAL SCORES INDICATE:

- Normally alert and vigilant
- No evidence of inattentiveness on this task
- Check Delay Task performance for level of impulsivity in self-paced condition
- Normal Delay Task performance would indicate ability to integrate feedback to stabilize responding
- Ability to suppress responding under conditions requiring sustained attention is in borderline range

BLOCK SCORES INDICATE:

- Performance deteriorates during session
- Behavioral suppression never reaches normal range

EDUCATIONAL PROGRAMMING

- Full psychological evaluation may be indicated

TASK MONITORING DATA

SEQ.	BLOCK: 1	2	3	TOTAL
X19	12	7	11	30
XX1	0	1	2	3
19X	2	6	7	15
XX9	1	0	0	1
X9X	0	1	1	2
X1X	1	2	2	5
XXX	3	8	9	20

This display shows you the sequence of digits that preceded each of the button presses. An X in this table indicates that any digit may have appeared. For example a button press registered in the X19 row is one which correctly followed the presentation of the "1" and "9" digits in sequence. A score in the "19X" row means the three digits preceding the button press could have been 1, 9, and 3 or 1, 9, and 7, etc. Instead of pressing the button right after the "1/9" presentation, the subject pressed the button after the next digit was displayed. The XXX refers to button presses that were unrelated to either a 9 or a 1. In the first block of the above example, the subject pressed the button correctly after 12 "1/9" presentations (X19), on two occasions one digit too late (19X), once when alerted by a 1 followed by a digit other than 9 (X1X) and three times after a digit that was unrelated to the appearance of a 1 or 9.

IV. FUNCTIONS MENU

The display you see after the PRINTER READY prompt, after the displays, or after the completion of any "function" is as follows:

START TASK OR PRESS A FUNCTION KEY
F1 IDENTIFY NEW SUBJECT
F2 ENTER RESULTS
F3 REDISPLAY CURRENT RESULTS
F4 LIST STORED SUBJECTS
F5 RECALL STORED RESULTS
F6 DELETE STORED RESULTS
F7 END SESSION

To perform any of these functions you simply press the appropriate function key, usually marked on the keyboard as F1, F2, etc. On those computers which do not have function keys, such as the Apple, you press the appropriate number key instead. For example if you wanted to identify a new subject, you would press the 1 key when the function name appeared. Functions F1-F3 and F7 will be described in this chapter. Those dealing with stored data, F4-F6, will be discussed in the next chapter.

F1 IDENTIFY NEW SUBJECT

The purpose of this function is to allow you to identify the subject before the task has been performed rather than afterwards. The choice is yours and does not affect how the data is displayed, stored or retrieved. The prompts and responses for this function have already been explained in **Chapter III, Section 3, Subject Identification.**

F2 ENTER RESULTS

For some computers, including the Apple, this function is necessary to let the GDS transfer data to the computer electronically. For all computers, this function gives you access to a new capability not previously described. It lets you manually enter the results of a GDS task into the computer, which can then interpret and store your data. You can use some of the DAP features even if you do not have your computer connected during the administration of the GDS.

When you press F2, the following prompts appear:

ELECTRONIC OR MANUAL DATA ENTRY (E/M)?>

If you respond E, the computer is ready and waiting for data to come from the GDS. (Most computers, except the Apple, are also ready to receive data whenever you see the functions menu.) If you change your mind about running a task, pressing RESET button ON THE GDS will return you to the function menu.

If you respond M and have not just identified a subject, the computer will show:

YOU MUST IDENTIFY A SUBJECT FIRST

SUBJECT NAME>

and proceed through the usual sequence of prompts to identify the subject. If you press <ENTER> without typing a subject name, the computer will proceed through the next set of prompts, but will not store your data.

Once the subject has been identified, the computer will ask:

DELAY OR VIGILANCE TASK (D/V)?>

to which you respond with D or V depending on the data you are entering.

If you previously stored a task for a subject and try to do so again, the computer will say:

SUBJECT ALREADY HAS THAT TASK STORED

You will need to re-enter the subject identification so the computer treats it as a new subject.

It then asks:

SAME PARAMETERS FOR ALL BLOCKS (Y/N)?>

If you ran the task with the same Interval and Block length for each block then answer Y. The computer will respond:

ENTER INTERVAL FOR BLOCK 1>

For the Delay Task, the standard interval is 6 (sec); for the Vigilance Task it is 1 (sec). The next prompt is:

ENTER LENGTH OF BLOCK 1>

The standard lengths for the Delay and Vigilance Tasks are 120 and 180 (sec), respectively. If you did not use the same parameters for all blocks, the computer will prompt you to enter the Intervals and Lengths for Blocks 2-4. Otherwise it assumes that the Block 1 Interval and Length you entered apply to the other blocks. (If you press <ENTER> after either the INTERVAL or LENGTH prompts, you will be returned to the functions menu without completing this routine.)

The computer will then display:

BLOCK:	1	2	3	4	TOTAL
INTERVAL	6	6	6	6	
LENGTH	120	120	120	120	480

ARE THESE CORRECT (Y/N)?>

If you respond N it will start over at the INTERVAL prompt.

Otherwise it will go on to ask:

DATA FOR SELECTOR POSITION 1>

Your response should be the number that you read from the back of the GDS when the Data Selector was at 1. The DAP will not let you press <ENTER> without typing a number and will not accept any entries greater than 99. After you respond, the computer will prompt you to enter the data for selector positions 2-10. When you finish that the computer will display the 10 numbers you entered plus appropriate sums or quotients, such as:

BLOCK:	1	2	3	4	TOTAL
REWARDS	11	12	10	13	46
RESPONSES	13	15	12	16	56
EFFIC RATIO	.85	.80	.83	.81	.82

ARE THESE CORRECT (Y/N)?>

Again you have the option of reentering the 10 numbers if you made a mistake. If you made no error, pressing Y starts the computer on the display routines to let you print the data and give you interpretations.

F3 REDISPLAY CURRENT RESULTS

The purpose of this function is to let you view the results again and to give you another opportunity to print them if you need to. The displays produced are exactly the same as those described in **Chapter III, Section 4, Data Display.**

F7 END SESSION

When you have finished using the computer, you should press F7. Failure to do so may result in the computer's not storing the last information it handled. If you press F7 and then decide you need to run the DAP again, pressing F7 will restart the DAP. (To restart the DAP on computers which do not have function keys, such as the Apple, you need to type RUN and press the RETURN key instead.)

309

V. STORED RESULTS

One of the valuable features of the DAP is that it stores the results obtained from administering the GDS. This enables you to print the results at a time and place separate from the task administration. It also allows you to collect sets of results to investigate one subject's responses over time, or differences among diagnostic groups. If you choose to use this feature, you will need to learn some additional steps to retrieve your data.

The data is stored in a floppy disk, magnetic bubble cartridge, or in computer memory, depending on the computer you are using. However, the operation of the functions has been designed to be the same regardless of the storage medium. The only difference is that the capacity to store results varies with the medium as is described in **Appendix B.**

The functions needed to manipulate stored results are F4-F6 on the function menu. The operation of each of them is described below. (Functions F1-F3 and F7 were described in the previous chapter.)

START TASK OR PRESS A FUNCTION KEY
F1 IDENTIFY NEW SUBJECT
F2 ENTER RESULTS
F3 REDISPLAY CURRENT RESULTS
F4 LIST STORED SUBJECTS
F5 RECALL STORED RESULTS
F6 DELETE STORED RESULTS
F7 END SESSION

F4 LIST STORED SUBJECTS

This function produces a list of the names of the subjects that you have stored. The names can be listed in either a quick or standard form as determined by your response to:

QUICK OR STANDARD LIST (Q/S)?>

If you respond Q, a typical list would be:

1 MICHAEL SMITH	D	V9
2 TAMMI JONES	D	
3 PUBLIC, JOHN Q.		V1

PRINT/CONTINUE/SKIP RESULTS (P/C/S)?

The "D" and "V" notations tell you which tasks are stored for each subject. For example, Michael has both tasks Delay and Vigilance "1/9") stored, but Tammi has only the Delay Task.

The last line, as always, gives you the option of printing the list. The numbers are assigned by the computer and are only temporary. If you delete a subject (see below), that subject's number is reassigned to the next one on the list and all the numbers from there down are changed.

If instead of Q you had responded S, you would see:

SUBJECT NUMBER 1

SUBJECT NAME:	MICHAEL SMITH
SEX (M/F):	M
BIRTH (MM/DD/YY):	05/09/76
EXAMINER NAME:	JOHN PSYCHE
COMMENT:	R/O ADD
TEST DATE (MM/DD/YY):	01/10/84

PRINT/CONTINUE/SKIP RESULTS (P/C/S)?

After you respond P or C to this prompt, the same information will be displayed for the next subject. If you answer S then you will be returned to the function menu. You can repeat this listing function as often as you like.

F5 RECALL STORED RESULTS

This is the function that allows you to bring back into the computer results which it filed away earlier. To do so you will have to tell the computer whose results you want and whether you want the Delay or Vigilance Task. The prompts are as follows:

ENTER SUBJECT NAME>

You should read **Chapter II, Section 6, Subject Name** to learn the optional ways you can answer this prompt. The computer will find the first subject whose name matches your response to the above prompt. It will display that subject's identifying information, as shown in the example under standard listing, except that the last line will be:

IS THIS THE CORRECT SUBJECT (Y/N)?>

If it is not the correct one and you answer N, the computer will find the next subject in the list whose name matches what you requested. If the computer does not find the match that you want, it will say so. If you have difficulty, use F4 LIST STORED SUBJECTS to help you find how you entered the name of the subject you are looking for.

If you answer Y, the computer will ask:

DELAY OR VIGILANCE TASK (D/V)?>

310

Answer D or V as needed. The computer will then start displaying the results of your subject's task in the same form as shown in **Chapter III, Section 4, Data Display.** When the results have all been displayed, you will be returned to the function menu. If you wish, you can redisplay the results using F3 REDISPLAY CURRENT RESULTS.

F6 DELETE STORED RESULTS

The storage capacity of your computer is limited (see **Appendix B**). In order to make room for storing more subjects' results, you will delete those subjects whose data you no longer need. You start the process by pressing F6, to which the computer responds:

ENTER SUBJECT NAME>

The format of what you enter is the same as that previously described. The computer finds the subject and asks:

IS THIS THE CORRECT SUBJECT (Y/N)?>

If you answer Y, you get one more chance to change your mind before the results are permanently erased. The computer asks:

DO YOU WANT TO DELETE THIS SUBJECT (Y/N)?

If you've changed your mind and answer N, you are returned to the function menu. If you answer Y, the computer will pause for a minute while it deletes the subject's test results. After it says:

SUBJECT DELETED

PRESS ENTER KEY TO CONTINUE

the function menu reappears.

The disk cartridge or tape we supplied included Michael Smith's data. You may want to try recalling it to see how these functions work.

311

APPENDIX A: System Requirements

The Gordon Diagnositic System (GDS)

The Data Analysis Program (DAP) is designed to simplify the collection of data generated by the GDS. Data is transmitted from the GDS to the computer via the standard communication link, an RS-232 connector. A female RS-232 connector is located on the back of the GDS. It "appears" to the computer like a serial printer, except that the data goes into the computer instead of out.

The RS-232 Cable

The GDS end of the cable needs to be a male 25 pin standard RS-232 connector. The requirement for the computer end depends on the computer you are using. Usually it is either a male or female 25 pin connector, but some computers use round connectors with fewer pins. As only 7 of the 25 pins are actually used, the cable itself also sometimes has fewer than 25 conductors.

Many computers will be able to run the DAP. The computer must have a serial communications port and should have another port to connect to a printer. The minimum memory requirement is 32K. The program and data storage can be on disk, recording tape, magnetic bubble cartridge, or Random Access Memory (RAM) depending on your computer. The DAP is written in Microsoft* or Applesoft* Basic, so you must have the appropriate one available on your machine. We cannot provide it with the DAP.

[*Microsoft and Applesoft are registered trademarks of Microsoft, Inc. and Apple Computers, Inc., respectively.]

The Printer

The printer is not essential to running the DAP, but a printer should be used to get the greatest benefit from the DAP. To provide graphic representations of the data on a variety of printers, the graphics are generated as a series of printed characters.

APPENDIX B: Specific Computers

Apple II

RS-232 port:	Super Serial Card & others
RS-232 cable:	25 pin, male/male
Storage:	floppy disk
Capacity (Subjects):	25
Startup:	automatic
Capital Letters:	automatic
Function Keys:	no
ENTER Key:	RETURN
BREAK Key:	CTRL & C
ESCAPE Key:	ESC

Sharp PC 5000

RS-232 port:	built in
RS-232 cable:	male/male
Storage:	magnetic bubble
Capacity (Subjects)	25
Startup:	automatic
Capital Letters:	Press CAPS key
Function Keys:	yes
ENTER Key:	arrow down, then left
BREAK Key:	ON/BRK
ESCAPE Key:	ESC

Radio Shack Model 100

RS-232 port:	built in
RS-232 cable:	25 pin, male/male
Storage:	tape cassette, computer memory
Capacity (Subjects):	6
Startup:	After loading from cassette, type RUN <ENTER>. Restarts automatically until program reloaded from tape.
Capital Letters:	press CAPS LOCK
Function Keys:	yes
ENTER Key:	ENTER
BREAK Key:	SHIFT & BREAK/PAUSE
ESCAPE Key:	ESC

IBM PC

RS-232 port:	Quadram and others
RS-232 cable:	25 pin, male/female
Storage:	floppy disk
Capacity (Subjects):	50
Startup:	automatic
Capital Letters:	press Caps Lock
Function Keys:	yes
ENTER Key:	arrow down, then left
BREAK Key:	Ctrl & Scroll Lock/Break
ESCAPE Key:	Esc

312

APPENDIX C: Organization of Data Files

This section describes the way the DAP stores its data. It is provided for the DAP user who wishes to write another computer program accessing the data.

Filename: TMONIT.DO
Type: Random Access (Serial in Model 100)
Record Length: 180 bytes
# of Records: 8 x maximum number of subjects stored
Record Format: Delay Task — 4 records

Records 1-4 for Blocks 1-4
Up to 60 three digit numbers/block each number is the tenths of a second from the previous button press

Vigilance Task — 4 records

Record 1 = 180 single digits for sequence of digits displayed
Records 2-4 for Blocks 1-3
180 spaces or hyphens each of whose positions corresponds to the character presented before a button press

In the Model 100, TMONIT.DO has been divided into DMONIT.DO and VMONIT.DO for the Delay and Vigilance tasks, respectively.

Filename: SUBJ.DO
Type: Random Access (Serial for Model 100)
Record Length: 220 bytes
# of Records: same as maximum number of subjects stored
Record Format:

CONTENT	bytes
Subject Identification	72
Subject Name	25
Space	1
Sex	1
Space	1
Birth	8
Examiner Name	13
Space	1
Comment	13
Space	1
Test Date	8
Delay Task Summary Statistics	72
1-10 Data Selector Displays each as 3 digits + space	40
Block Intervals each as 2 spaces + 2 digits	16
Block Lengths each as space + 3 digits	16
Delay Task Identifier = D	2
Vigilance Task Summary Statistics Same format as Delay Task	72
Vigilance Task Identifier = V9 or V1	2

APPENDIX D: Making Copies

All microcomputers and their storage media are subject to occasional failure. As a result, it is virtually essential to make a "backup" copy of anything important. The backup usually has the same form as your storage medium. For example, if you are using floppy disks, your backup will be a floppy disk which contains the exact same information as your original.

To provide for this at the outset, you should follow the directions supplied with your computer manual to copy the original DAP and put it away in a safe place. Use the copy to operate your computer. If the operating one fails, recopy the original DAP onto the cassette, disk, or cartridge you use every day.

If you are collecting a large amount of data on your storage medium, the same warning about making a backup copy is applicable. You may want to purchase additional cassettes, disks or cartridges to make backup copies of your stored data. All three storage media can be reused repeatedly when you no longer need the data you have saved on them.

ADDITIONAL INFORMATION:

The Gordon Diagnostic System may be used for the clinical assessment of stimulant drug response in the treatment of children with Attention Deficit Disorder (ADD). With the next publication of the instruction manuals, additional instructions (draft attached) will be provided for its use in this regard.

The foundation for the use of the GDS data used in this regard is based on research work performed by Dr. Russell Barkley and Dr. Mark Rapport and presented to the American Psychological Association Meeting in August 1985.

Russell Barkley presented data on a multimethod clinical protocol for the clinical assessment of stimulant drug responding. Barkley's approach involves a triple-blind, placebo controlled crossover design using two doses of methylphenidate (0.3 mg/kg and 0.5 mg/kg) and a placebo. To assess drug response, he uses parent and teacher ratings of ADD behaviors, problem settings and drug side effects; the GDS Tasks; and clinic playroom observations of ADD behavior during academic performance. Most of the measures in the protocol were sensitive to the effects of stimulant medication at both doses. The Commission score was found sensitive at the higher dose. Barkley urged that this objective, multimethod assessment protocol replace the current practice of relying upon subjective impression.

Mark Rapport from the University of Rhode Island presented data from an extensive drug response study that incorporated school-based observations of behavior. Analyzed on a group basis, most of the variables in the protocol (i.e. percent on task, percent work complete, percent work correct, teacher rating scales of attention, GDS Vigilance Task Commission Score) there were significant linear relationships with dosage levels of medication. There were also significant differences on these measures between each level of medication. In the analyses of group trends, the Vigilance Task Ommission score did not reach significance as a variable responsive to medication (the Delay Task was not used in this study). Subject-by-subject analyses, however, indicated that the failure of the score to reach significance was related to the idosyncratic response of children to the various dosage levels. In other words, a particular child might have improved his score significantly at one dose, but not at another - a phenomenon very much present, but to a lesser degree for the other variables. Rapport viewed these data as pointing up the need to judge children's drug response on an individual basis using a comprehensive battery of tests.

MONITORING OF PHARMACOTHERAPY

(DRAFT)

The Gordon Diagnostic System (GDS) can be used as part of an evaluation battery for the monitoring of pharmacotherapy. While the GDS is useful in this process, it is important that decisions regarding dosage levels include a comprehensive series of other assessment techniques.

Monitoring of pharmacotherapy is done optimally by testing the child when s/he is receiving several dosages of medication or a placebo administered in a random order. Only the pharmacist should know whether a child is on placebo or one of the dosage levels of medication. The following procedure is suggested:

- o The child should be administered a dosage (placebo or otherwise). After 5-7 days the child should be tested with the GDS tasks.
- o Following the GDS testing, the next dosage should be administered and, 5-7 days later, the GDS testing should be repeated.
- o This procedure should be repeated for each drug condition.
- o At the end of the trials, scores for each of the GDS tasks should be tabulated. A graph can be constructed which presents the GDS scores at each dosage level.

INTEPRETATION OF GDS SCORES FOR ANALYSIS OF A DRUG EFFECT

Research to date suggests that the Total Commission score is the most sensitive to the effects of stimulant medication. As such, improvement in this score from placebo condition to drug condition would indicate improvement in a child's ability to sustain attention. A general rule of thumb is to consider a score improved if the child's score changes from one range to another according to the Threshold Tables. For example, a child's score is considered significantly improved if, for example, the score moves from the Abnormal Range to the Borderline Range (or Borderline to Normal). The effects of various dosage levels can also be gauged with this method.

While current studies have been most supportive of the Commission score, it is suggested that all the scores be analyzed for changes related to medication status. It is again important to emphasize that decisions surrounding the monitoring of pharmacotherapy should be based only on a comprehensive set of clinical data.

Journal of Abnormal Child Psychology, Vol. 7, No. 3, 1979, pp. 317-326

The Assessment of Impulsivity and Mediating Behaviors in Hyperactive and Nonhyperactive Boys[1]

Michael Gordon[2]
The Ohio State University

Twenty boys (6-8 years) rated by their teachers as hyperactive and a matched sample of nonhyperactive boys performed a task that required them to withhold responding for a set time interval in order to be rewarded (DRL 6-second schedule). Half of each group worked on a one-button console while the other half was provided with additional collateral buttons. Results indicated that hyperactive children were relatively unable to perform efficiently on the task, and that this deficit endured regardless of age, IQ, or experimental condition. DRL was thus found to discriminate accurately between teacher-rated and parent-rated hyperactive and nonhyperactive children. Furthermore, a wide variety of self-generated mediating behaviors was observed, and it was determined that a child's DRL performance was related to the kind of mediating behaviors he displayed. Results are discussed in terms of the clinical assessment of hyperactivity and the training of impulsive children.

Although the traditional approach has been to regard hyperactivity as a disorder of activity level, there is little evidence that the majority of children diagnosed as hyperactive are any more active, on the average, than a matched sample of nonhyperactive children, regardless of how activity is measured (Davis, Sprague, & Werry, 1969; Doubros & Daniels, 1966; Hutt, Jackson, & Level, 1966; Johnson, 1972; Lee & Hutt, 1964; McConnel, Cromwell, & Bialer, 1964; Ounsted, 1955; Schulman & Reisman, 1959). Moreover, viewing hyperkinesis solely in terms

Manuscript received in final form December 15, 1978

[1] This report is based upon a dissertation submitted by the author in partial fulfillment for the Ph.D. degree. The efforts of Malcolm Helper, David Hothersall, Donald R. Meyer, Marilyn Moody, and Charles Wenar are gratefully acknowledged. The investigation was conducted while the author was a National Research Fellow, supported by Public Health Service Training Grant No. 5 T32 MH14608-02 to The Ohio State University.

[2] Address all correspondence to Michael Gordon, Upstate Medical Center, Division of Child and Adolescent Psychiatry, 750 East Adams Street, Syracuse, New York 13210.

316

of activity makes it difficult to understand many of the other symptoms that are often associated with this disorder, e.g., antisocial behavior, learning disabilities, distractibility, and excitability. The disenchantment with focusing on the quantity of a hyperactive child's activity has led some investigators to conceptualize hyperkinesis as a disturbance of impulse control and sustained attention. According to Douglas (1972), what typifies the hyperactive child is his inability to "stop, look and listen." She reviewed a body of research that appears to suggest that what characterizes the hyperactive child's behavior is not necessarily the sheer amount of activity, but its poor organization and control.

While the supporting studies cited by Douglas (1972) are intriguing, there is still need for a more straightforward assessment of the hyperactive's impulsivity. There is a tested technique, derived from research in operant conditioning, that provides an excellent opportunity to measure the hyperactive child's ability to inhibit behavioral responding. In this procedure, termed Differential Reinforcement of Low Rate Responding (DRL), responses that occur before a set time interval has passed are not reinforced and, moreover, they serve to reset the timer governing reinforcement. Thus a subject performing on a DRL 6-second schedule must wait at least 6 seconds between responses in order to receive a reinforcement. If a response is made before the 6 seconds have elapsed, the timer will reset and another 6 seconds must pass before a successful response can be made.

Normal laboratory animals performing on this task are able to develop efficient response patterns and typically come to be reinforced for 70% to 80% of the responses they make (Hothersall, Alexander, & Slonaker, 1972). As for human DRL performance, Weisberg and Fragakis (1967) demonstrated that even young children, rating from 15 to 40 months, were able to maintain low levels of responding. Latency-age boys (Stein & Landis, 1973) and college students (Randolph, 1964) have also been shown to perform efficiently, although the samples for most studies were extremely small. One major question posed by the present study is: How would a hyperactive child fare in a DRL situation in which behavioral responding must be suppressed? Might the DRL procedure provide an objective measure of impulsivity?

Researchers investigating DRL performances have also observed that subjects in a time-based situation will often engage in a sequence or chain of behaviors between reinforceable responses. The notion is that this collateral behavior sequence helps the subject "wait out" the required temporal delay between responses. Examples of such behaviors in animals include wood chewing and tail licking (Slonaker & Hothersall, 1972). Similarly, human subjects have become convinced that these extra mediating behaviors may set up the reinforced responses (Bruner & Revusky, 1961). Experimentally precluding such collateral behaviors has led to a disruption of the subject's ability to make temporal discriminations (Stein & Landis, 1973; Stein & Flanagan, 1974). The

frequency of extra movements made by hyperactive children during a delayed reaction time task was positively related to the efficiency of their performances (Douglas, 1972; Cohen, cited in Douglas, 1972), suggesting that collateral behaviors may play a role in helping these children gain impulse control. Another intent of the current study was to examine whether or not a situation that facilitates the development of collateral behaviors helps to improve the DRL performance of hyperactive and, for that matter, normal children. And, regardless of the availability of experimentally introduced opportunities for collateral responding, will these children spontaneously elicit mediating responses?

METHOD

Subjects

Twenty hyperactive boys between 6 and 8 years of age (72-97 months) and a matched sample of nonhyperactive boys participated in the study. The subjects were selected from the client roster of a child guidance clinic; they were all nonretarded and enrolled in school at an age-appropriate grade level. A child was considered hyperactive if his teacher rated him as such on the hyperactivity factor of the Behavior Rating Scale (Conners, 1969). In accordance with the findings of a normative study (Werry, Sprague, & Cohen, 1975), those children with a total score of at least 15 on this scale were allotted to the hyperactive group, while subjects receiving a score of less than 15 made up the control group. The two groups were statistically comparable for age.

Design

The experiment essentially followed the design and methodology presented by Stein and Landis (1975). Four equal-sized groups of 10 subjects each were involved: two groups of hyperactive and two groups of nonhyperactive boys. One hyperactive and one nonhyperactive group performed on a DRL 6-second schedule without the availability of collateral buttons (DRL ONLY), while the other two groups worked on DRL using the five-button console (DRL + COLLATERAL).

Apparatus

Each subject was seated before a console containing either one (DRL ONLY) or five (DRL + COLLATERAL) microswitch buttons mounted 4.0 cm

318

apart. On a separate console, positioned immediately behind the button box, was a chassis containing a large six-digit add counter, which indicated the cumulative number of reinforcements (points) earned, and a red light, which lit for 2 seconds upon completion of a successful response. For the DRL ONLY condition, the one red button was programmed to produce reinforcement according to a DRL 6-second schedule. In the DRL + COLLATERAL condition, only one red button was programmed with the DRL schedule. The remaining four black collateral keys did not yield reinforcements, but they were individually monitored for the number of times they were pressed and the response pattern was retained by an Esterline Angus Event Recorder. All sessions were videotaped in their entirety. The programming equipment, counters, timers, recorders, and monitoring equipment were housed in a separate room, which was joined to the experimental room by a one-way mirror and an intercom system.

Procedure

Prior to the testing session, the mother of every subject received an appropriate introduction to the study. The child was then taken individually to the experimental room, which contained only a table, chairs, and the testing apparatus. He then received the following instructions (adapted from Stein & Landis, 1975).

> You're going to play a game in which you will get a chance to win a lot of M&M's. Do you see this light [pointing to the red indicator on the reinforcement box]? Every time you make this light go on you will earn an M&M and this counter [pointing] will keep track of how many M&M's you've won. At the end of the game I'll give you all the M&M's that you earned. Now, to make the light go on all you have to do is push this red button [pointing], and wait a little while before pressing it again. You just press this red button, wait a while, then press it again. If you press it again too soon, though, then you will not get a point, the light won't go on, and you'll have to wait a while before you can press it to get another point. But if you push the button, wait a while, then push it again, you'll get a point every time.

For the DRL + COLLATERAL groups, these instructions were added:

> These four black buttons will not give you points if you press them but, if you want, you can play with them while you're waiting to press the red button again.

The general instructions continued:

> Do you have any questions? [All questions were answered.] OK, I'll leave the room now and turn on the machine. When you hear my voice coming over the speaker saying, "You can start now," then you can begin playing the game. I'll come back when the game is over and bring you your M&M's. Have fun!

Each subject had a total of 15 minutes to earn points on the DRL schedule. At the session's end he was asked, "How did you do it?" and "Did you do

319

anything to help you wait?" Following his response, he was paid his due of M&M's, thanked for his participation, and returned to his mother

Measures

The DRL task yields three measures of a subject's performance the total number of responses, the number of reinforcements earned, and the efficiency score. Since this last score represents the percentage of reinforced responses, i.e., the number of reinforcements divided by the number of responses, it is considered to be the most sensitive indicator of a child's performance. Every subject's response, reinforcement, and efficiency scores were obtained for each of three 5-minute time blocks, and then totaled for the entire 15-minute session. Also monitored were the number of responses emitted on the collateral keys (if available) and the pattern in which they were pressed. Data on IQ, exact age, school grade, and parent ratings were also collected.

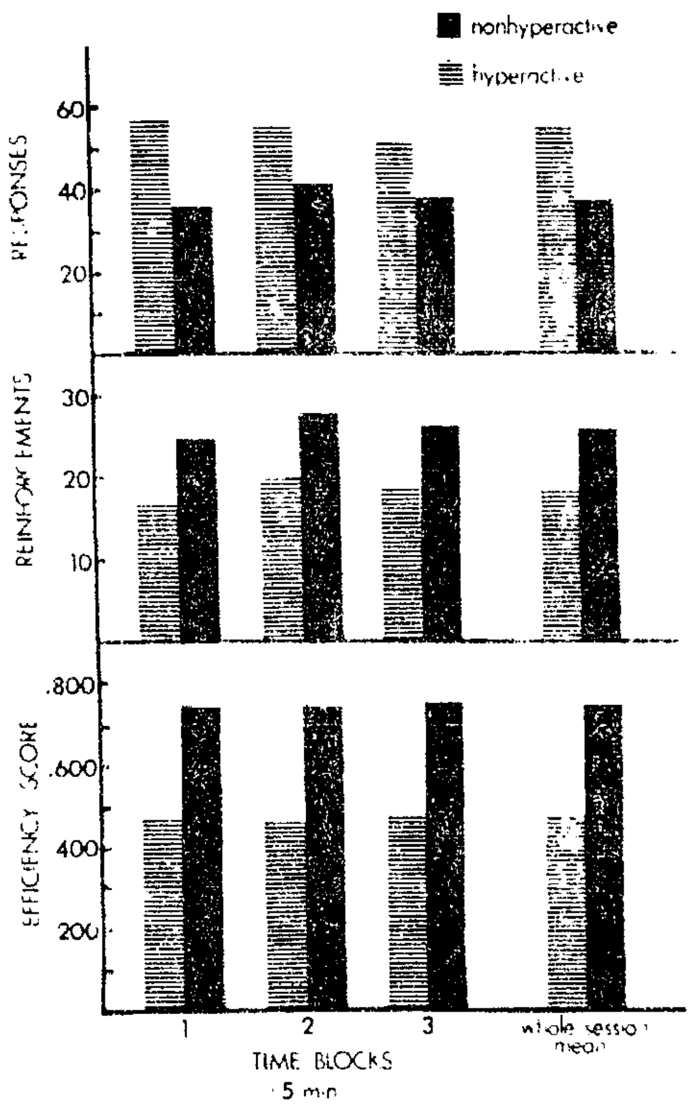
RESULTS

DRL Analysis

Analyses of variance (ANOVAs) were performed on the three operant measures, and the mean scores for each trial and for the whole session are shown in Figure 1. There was a highly significant main group (hyperactive vs. nonhyperactive) effect for all three DRL variables, especially the efficiency score, $F(1, 36) = 18.33$, $p < .0013$. Children rated as hyperactive by their teachers tended to emit more responses and receive fewer reinforcements during each of the time blocks than did children judged nonhyperactive. This main effect remains strong if the data are reanalyzed according to the parents' classifications of hyperactivity. The relative inefficiency of these hyperactive children suggests that they are significantly more impulsive as measured by the DRL 6-second schedule.

There were no Condition (DRL ONLY vs. DRL + COLLATERAL) effects on any of the dependent variables (efficiency score, $F(1, 36) = 133$, $p < .717$), so that whether a child played with the one-button or the five-button console had an insignificant effect upon his DRL performance. A few Group X Condition interactions did emerge. Hyperactive children in the DRL + COLLATERAL condition tended to make more responses than did their counterparts in the DRL ONLY condition, $F(1, 36) = 6.4$, $p < .016$, and nonhyperactive COLLATERAL subjects did significantly better according to the efficiency score than did the nonhyperactives in the DRL ONLY situation.

The ANOVAs yielded only one marginally significant main effect for Trials (time blocks): All groups tended to receive more rewards for their second-



nonhyperactive
hyperactive
RESPONSES
60
40
20
REINFORCEMENTS
30
20
10
EFFICIENCY SCORE
.800
.600
400
200
1
2
3
whole session mean
TIME BLOCKS
5 min

Fig. 1 [illegible]

321

trial performance, $F(2, 72) = 3.08$, $p < .051$ In general, the second trial was the most characteristic of a child's DRL work. For nearly all the various analyses, what held true for each trial taken separately, or for the session as a whole, usually held true for the second-trial measures and, again, particularly for the second time block efficiency score. This observation is supported by the intertrial correlations for each measure Since the correlations between Trial 2 and Trial 3 tend to be quite high, mean $r = .69$, $p < .001$, it would suggest that a boy's performance was apt to stabilize by the second trial When these interblock correlations are computed for each group separately there is evidence that the DRL conduct of nonhyperactive subjects generally consolidated from one trial [illegible] next [illegible] they locked into a relatively efficient pattern, mean [illegible] $p <$ [illegible] The hyperactive children were more erratic across trials than the nonhyperactive children, mean $r =$ [illegible]4, $p <$ [illegible], although their performance was also relatively stable as the interblock correlations indicate

In order to uncover relationships between the operant scores and a variety of descriptive data all variables generated by this study were intercorrelated It was first found that the subject's age is thoroughly unrelated to any other aspect of his performance or to the hyperactivity ratings supplied by the teacher or parent The teacher ratings of hyperactivity using the Behavior Rating Scale (Conners, 1969) were significantly correlated with all DRL variables, their correlation with the whole session efficiency score, for example, was $r =$ [illegible]3, $p <$ .01 The parent rating was also tied to the major DRL scores (mean $r =$ [illegible], $p <$ [illegible]), though slightly less than were the teacher ratings The correlation between parents and teachers themselves was $r =$ [illegible]1 highly significant ($p <$ .001) but not so high as to instill total confidence in their reliability Parents and teachers did closely agree (34/40 or [illegible] of the cases) on the gross classification of a child as hyperactive or nonhyperactive Of the six contesting cases, parents rated four teacher-classified nonhyperactives as hyperactive and two teacher classified hyperactives as nonhyperactive It seems therefore that parents were apt to classify more children as hyperactive Yet the teachers mean ratings of the hyperactive boys (2[illegible]) was slightly higher than those rendered by parents (2[illegible]) So teachers classified fewer children as hyperactive, but they judged the hyperactives as more severely hyperactive

Analysis of Mediating Behaviors

As has been the case for almost all animal and human subjects working on a DRL schedule, the children who participated in this study tended to engage in a variety of interresponse behaviors that were either physically manifested or were reported to the examiner in response to a standard, end-of-session question "How did you do it?" or "Did you do anything to help you wait?" A list of some of the mediating behaviors that were observed appears in Table

322

Table I. A List of Observed Mediating Behaviors

Circling DRL button with finger 9 times
Swinging legs 10, 12, or 20 times
Counting with lips
Counting out loud — numbers or ABC's
Blowing on reward box
Singing out loud
Shaking reward box 10 times
Hitting knee with right hand 20 times
Foot tapping 16 times
Tapping finger 10 times on button box
"Walking" fingers around DRL button 9 times
Stomping with foot 9 or 10 times
Running around table 1 time
Hitting side of box
Jumping jacks 4 times
Hitting collateral buttons

I. (It should be noted that only 5 of the 20 Collateral group used the collateral buttons for their interresponse regimens.) In order to be considered a mediating behavior, the sequence had to be repeated among at least 10 responses within a particular trial. Judgments concerning collateral responding were made blind from the videotapes. As it turned out, every subject engaged in, or reported, some mediating behavior. Ninety percent of the hyperactive subjects engaged in an observable mediating behavior, compared to 45% of the nonhyperactives. On the other hand, just 30% of the hyperactive boys reported the use of a nonbehavioral mediator, as opposed to 80% of the nonhyperactive subjects. A roughly comparable percentage of each group used both approaches (hyper-actives = 20%, nonhyperactives = 25%).

There was an inverse correlation between the use of a physical collateral and both the efficiency ($r = .32$, $p < .05$) and reinforcement ($r = .55$, $p < .005$) scores, especially after the first trial. Hence, children who used an observable collateral scheme did worse on the DRL schedule; they were also apt to be rated as more hyperactive by their parents and teachers. The opposite pattern emerged for cognitive mediators as there was a very high correlation between the children's reports of counting and all measures for each trial, particularly blocks 2 and 3, mean $r = .64$, $p < .001$. Thus subjects who engaged in cognitive collaterals fared better on DRL and were usually judged less hyperactive than their noncounting counterparts. Boys who used both strategies during the session also had better scores, though not as high as the scores of those who only counted.

DISCUSSION

The prediction that children classified as hyperactive would encounter greater difficulties in managing the DRL task than would nonhyperactives was confirmed by the results of this study. According to every indicator of DRL performance, the hyperactive boys were relatively unable to refrain from making a high percentage of incorrect responses. Consequently, these results lend further credence to Douglas's (1972) proposal of an impulsivity dimension underlying the hyperkinetic syndrome and to the validity of other measures of impulsivity presented by Sykes (cited in Douglas, 1972) and Campbell (cited in Douglas, 1972).

The experimental Condition (DRL ONLY vs. DRL + COLLATERAL) failed to emerge as a significant source of variance. However, given the observation that children engaged in many self-generated interresponse behaviors, this finding is not at all surprising. The fact is that a child will develop some type of interresponse strategy whether or not the experimenter provides a few extra buttons or telegraph keys. Indeed, in their attempts to quantify collateral responding, it appears that past investigators have missed the rich bounty of self-elicited behaviors that children exhibit during the course of a DRL session. Aside from the diversity, these schedule-induced behaviors are fascinating for the rigid manner in which they are executed; once a child chanced upon a successful strategy, he usually stuck with it ardently.

That the proclivity to engage in a physical or cognitive mediating behavior was so closely related to DRL performance as well as to ratings of hyperactivity is a highly intriguing result. Even with a relatively gross coding system, a strong pattern surfaced in which boys who displayed physical collaterals did worse on DRL and were rated more hyperactive than were boys who used cognitive means. In addition, many more hyperactive than nonhyperactive boys generated physical collaterals, an observation that may support the notion that a certain segment of the hyperactive child's hyperactivity is actually a series of mediating behaviors developed in an effort to control his impulses.

While it appears certain that cognitive mediators are more efficient strategies than physical ones, it is impossible to determine from these data whether physical collaterals actually hindered DRL efficiency or simply failed to facilitate it as much as the cognitive types did. The only hint available is the finding that the 18 subjects who employed physical collaterals during several trials had significantly higher efficiency scores than did the 9 children whose physical mediating was limited to one time block, $p < .05$. The precise nature of the interaction between a physical collateral and the DRL performance of hyperactive children is a very important avenue of research that warrants further exploration.

324

An unexpected aspect of these results was the extent to which the DRL variables could accurately discriminate between groups. According to the inter-trial correlations, the efficiency score can serve as a reliable and stable predictor of teacher-classified hyperactivity. Consequently, it is possible that this task could be developed into a useful clinical tool that could aid the diagnostician in the assessment of impulsivity and the classification of hyperactivity.

REFERENCES

Bruner, A., & Revusky, S. A. Collateral behavior in humans. *Journal of the Experimental Analysis of Behavior* 1961, *4*, 349-350.

Cantwell, D. P. *The hyperactive child. Diagnosis, management, current research*. New York: Spectrum Publications, 1975.

Conners, C. K. A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry*, 1969, *126*, 152-156.

Davis, K., Sprague, R., & Werry, J. Stereotyped behavior and activity level in severe retardates. The effects of drugs. *American Journal of Mental Deficiency*, 1969, *72*, 721-727.

Doubros, S., & Daniels, G. An experimental approach to the reduction of overactive behavior. *Behaviour Research and Therapy*, 1966, *4*, 251-258.

Douglas, V. I. Stop, look and listen. The problem of sustained attention and impulse control in hyperactive and normal children. *Canadian Journal of Behavioral Science*, 1972, *4*, 259-282.

Gittis, A., & Hothersall, D. DRL performance of juvenile rats with septal lesions. *Physiological Psychology*, 1974, *2*, 38-42.

Hothersall, D., Alexander, D., & Slonaker, R. The DRL deficit of rats with septal lesions. Effects of extended training in a mediated environment. *Psychonomic Science*, 1972, *29*(1), 34-36.

Hutt, D., Jackson, P., & Level, M. Behavior parameters and drug effects. A study of a hyperkinetic epileptic child. *Epilepsia* 1966, *7*, 250-259.

Johnson, C. F. Limits on the measurement of activity levels in children using ultrasound and photoelectric cells. *American Journal of Mental Deficiency* 1972, *77*, 301-310.

Lee, D., & Hutt, C. A play-room designed for filming children: A note. *Journal of Child Psychology and Psychiatry*, 1964, *5*, 263-265.

McConnel, T. B., Cromwell, R. L., & Bialer, I. A. Studies in activity level. VIII. *American Journal of Mental Deficiency*, 1964, *68*, 647-651.

Ounsted, C. The hyperactive syndrome in epileptic children. *Lancet*, 1955, *269*, 303-311.

Randolph, J. J. A further examination of collateral behavior in humans. *Psychonomic Science*, 1964, *3*, 227-228.

Schulman, J. L., & Reisman, J. M. An objective measure of hyperactivity. *American Journal of Mental Deficiency*, 1959, *64*, 455-456.

Slonaker, R. L., & Hothersall, D. Collateral behaviors and the DRL deficit of rats with septal lesions. *Journal of Comparative and Physiological Psychology*, 1972, *80*, 91-96.

Stein, N., & Flanagan, S. Human DRL performance, collateral behavior and verbalization of the reinforcement contingency. *Bulletin of the Psychonomic Society*, 1974, *3*, 27-28.

Stein, N., & Landis, R. Mediating role of human collateral behavior during a spaced responding schedule of reinforcement. *Journal of Experimental Psychology*, 1973, *97*, 28-33.

Stein, N., & Landis, R. Differential reinforcement of low rates performance by impulsive and reflective children. *Journal of Experimental Child Psychology* 1975, *19*, 37-50.

Weisberg, P., & Fragakis, C. J. Analyses of DRL behavior in young children. *Psychological Report*, 1967, *21*, 709-715.

Werry, J. S., Sprague, R. L., & Cohen, M. N. Conners' Teacher Rating Scale for use in drug studies with children. An empirical study. *Journal of Abnormal Child Psychology* 1975, *3*(3), 217-229.

325

The Objective Assessment of Attention Deficit Disorders

Michael Gordon, Ph.D.
State University of New York, Upstate Medical Center

F. Daniel McClure, Ph.D
David C. Wilson Neuropsychiatric Hospital

Objective and reliable techniques for the assessment of Attention Deficit Disorders (ADD) with and without hyperactivity in children have remained largely unavailable to psychologists, educators and pediatricians. As a consequence, they have tended to base their evaluations of ADD solely on observations and teacher reports, or on measures such as the Wechsler scales, which have been shown to discriminate poorly between groups of hyperactive and nonhyperactive children. Over the past 6 years we have been developing a behavioral measure of ADD, called the Gordon Diagnostic System (GDS), which has shown considerable promise as a precise, valid and efficient technique for the diagnosis of attention disorders. The GDS contains two measures: the Delay Task, which measures the ability to inhibit responding; and the Vigilance Task, which assesses sustained attention. In our research studies, we have shown that these game-like tasks differentiate accurately between hyperactive and nonhyperactive children. Normative data on GDS performance have been established based on a study of 220 nonhyperactive boys and girls from 6 through 11 years of age. Information on test-retest reliability is also available.

INTRODUCTION

Clinicians, educators, and physicians have been in chronic need of reliable and valid techniques for the assessment of Attention Deficit Disorders with and without hyperactivity (ADD) and impulsivity in children. The instruments that are currently available for the diagnosis of ADD fall into three basic categories: those actually intended for the measurement of other aspects of behavior which have been adapted to assess impulsivity, those which are largely limited to laboratory investigations, and questionnaires.

Paper presented at the 91st Annual Convention of the American Psychological Association at Anaheim, CA. August, 1983

Correspondence address 750 East Adams Street
Syracuse, New York 13210

326

Most prominent among the instruments originally designed for the clinical assessment of other behavioral domains is the Wechsler Intelligence Scale for Children-Revised (WISC-R). While selected subtests or groups of subtests from the WISC-R are widely used in the identification of hyperactivity, studies have shown that they do not discriminate accurately between groups of hyperactive and nonhyperactive children (Douglas, 1972). Moreover, those subtests assess a wide range of abilities and traits that are largely unrelated to impulsivity (Douglas, 1972). Thus, while all children with ADD may perform poorly on these measures, there are many nonhyperactive children (e.g. learning disabled, or those with visual-motor deficits) who may also have difficulties with them for reasons other than impulsivity

Among those instruments best suited for laboratory investigations, the most effective in differentiating hyperactive from nonhyperactive children are the Continuous Performance Test (CPT: Rosvold & Mirsky, 1956), Delayed Reaction Time Test (DRTT; Cohen, Douglas & Morganstern, 1972), and the Matching Familiar Figures Test (Kagan, 1966). While the CPT and DRTT have been demonstrated to be excellent diagnostic measures, their technological sophistication and prohibitive expense make them unsuitable for widespread clinical use. The Matching Familiar Figures Test is quick, easy to administer, and has been demonstrated to be diagnostically sound. However, its reliability and validity have been the subject of considerable controversy which persists despite revisions of the measure (Egeland & Weinberg, 1976).

There are a variety of other measures available such as the Gardner Steadiness Tester (Gardner, 1979), Human Figure Drawings (Machover, 1949), certain configurations of Rorschach indices, and paper-and-pencil "vigilance" tasks. As a group, these tend to suffer drawbacks stemming from impracticality of administration, limited yield of information, or inadequate standardization.

Behavioral inventories or questionnaires have also enjoyed widespread use, and many, such as the Teacher Rating Scale (Conners, 1969), the Child Behavior Problem Check List (Achenbach, 1978), and the Self-Control Rating Scale (Kendall & Wilcox, 1979), have demonstrated good discriminative validity. However, there are serious inherent drawbacks to such a format, including rater bias, the possibility of "halo" effects, item overlap, vague rating criteria, issues of basic validity, and the fact that questionnaires yield no direct behavioral data.

There is thus an obvious and compelling need for a more straightforward, clinically applicable, and precise measure of impulsivity in the child with ADD. The Gordon Diagnostic System (GDS) shows considerable promise as a device which can help fill this gap in clinical service delivery. The GDS is an easily-administered, game-like task which measures levels of impulsivity without interference from other factors such as intelligence or visual-motor skills. It provides objective data on a particular child's ability to inhibit behavioral responding. Through observation of the child's performance and from responses to several post-test inquiries, important insights can be gained into a child's problem-solving style. This information can have strong bearing on treatment recommendations and educational planning. Thus, the GDS is unique because it yields not only quantitative data, but also a wealth of pertinent behavioral information which is crucial to the diagnostic process.

DESCRIPTION OF THE GDS

The GDS is a portable, electronic unit which allows for the administration of two tasks. The Delay Task is based upon a Differential Reinforcement of Low Rates (DRL) behavior schedule and requires the child to inhibit responding in order to gain a reward. Specifically, the child is instructed to press a button, wait, and then press the button again. If he/she refrains from responding for at least six seconds, a light flashes and a reward counter increments. If the child responds before the interval elapses then the timer resets and no reward points are recorded. The GDS Delay Task generates three major scores: the number of responses (button presses), the number of correct responses (Rewards) and the Efficiency Ratio which represents the percentage of correct responses.

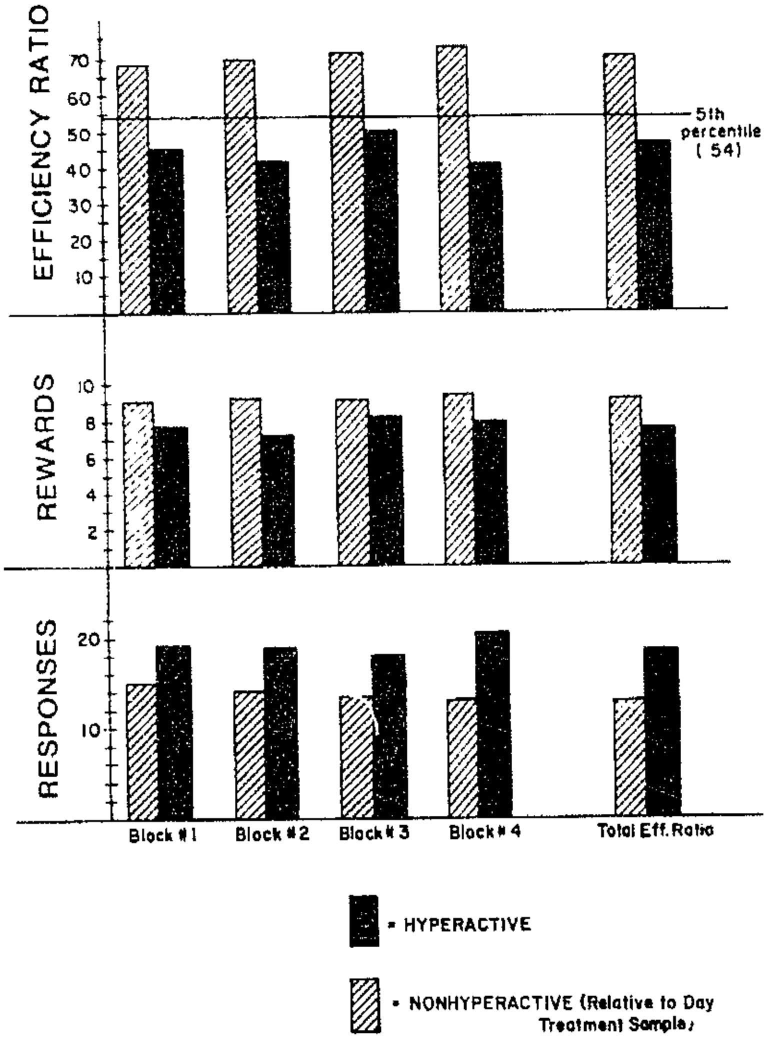
The GDS also contains the Vigilance Task which is a version of the Continuous Performance Test (Rosvold, et al., 1956). The child is presented with a series of digits on the front display and is told to press the button every time a "9" appears that had been preceded by a "1". The GDS records the number of correct presses, the number of times the child fails to press the button when the "1/9" combination appears (errors of Omission), and the number of extraneous button presses (errors of Commission). For testing younger children the GDS contains a "1" mode, which requires the subject to press the button every time a "1" appears. The same performance measures are recorded.

RESEARCH STUDIES WITH THE GDS

Data concerning the GDS Delay Task come from six studies conducted over as many years. The original experiment (Gordon, 1979), performed on an early version of the GDS, compared 20 hyperactive and 20 nonhyperactive children selected from the patient roster of a child guidance clinic. There were highly significant differences between the two groups for all three Delay Task variables: Responses, Rewards and the Efficiency Ratio (Rewards/Responses). The children classified as hyperactive had profound difficulties in dealing successfully with the delay inherent in the task.

The original finding was recently replicated (McClure & Gordon, in press) with patients placed in a day treatment program for seriously disturbed youngsters. Differences in Delay Task efficiency ratios between the 20 children classified as hyperactive and the 20 nonhyperactive ones in the comparison group were strikingly significant ($p < .0001$).

DELAY TASK PERFORMANCE OF HYPERACTIVE VS. NONHYPERACTIVE SUBJECTS: DAY TREATMENT SAMPLE



To further judge the discriminative ability of the GDS *post hoc* analyses were performed with groups constituted on the basis of being identified as hyperactive or nonhyperactive by four criterion measures These included the Teacher Rating Scale (Conners, 1969), the Self-Control Rating Scale (Kendall & Wilcox, 1979), the Teacher Report Form (Achenbach, 1980), and the Matching Familiar Figures Test - 20 (Cairns & Cammock, 1978) For subjects who were identified as hyperactive (n=8) or nonhyperactive (n=8) on the basis of all four criteria, there was agreement between classifications derived from the GDS Delay Task efficiency ratio and the criterion measures in 15 cases (94%, Fisher's Exact Test, p=.0007). When groups were classified as hyperactive or nonhyperactive by at least three of the four criterion measures, there was agreement in 29 of 32 cases (91%, Fisher's Exact Test, p= 000003)

DIAGNOSTIC ACCURACY OF THE DELAY TASK. DAY TREATMENT SAMPLE

Criterion Measures. Teacher Rating Scale (Connors, 1969)
Self-Control Rating Scale (Kendall & Wilcox, 1979)
Teacher Report Form (Achenbach, 1980)
Matching Familiar Figures Test (Cairns & Cammock. 1978)

Hyperactive Group: Identified as Hyperactive on at least 3 measures

Nonhyperactive Group. Identified as Nonhyperactive on at least 3 measures

	Efficiency Ratio High (> .55)	Efficiency Ratio Low (≤ .55)
Hyperactive (n = 15)	1	14
Nonhyperactive (n = 17)	15	2

p = .000003 (Fisher Exact Test)

There was agreement between the Delay Task Efficiency Ratio and the criteria on 91% of the cases.

Two additional studies appear to establish the validity of the GDS as a measure of impulsive/hyperactive behavior. In a group of 31 nonhyperactive subjects, GDS performance was significantly correlated with factors derived from teacher ratings. It was also related to impulsivity as measured by the

Bender-Gestalt test and the Rorschach Inkblots. Analyses of data from the files of 32 nonhyperactive outpatient children also indicate strong relationships between GDS performance and parent ratings of impulsive behavior.

CORRELATIONS BETWEEN DELAY TASK TOTAL EFFICIENCY RATIO AND COLLATERAL MEASURES

Day Treatment Sample-Combined Hyperactive and Nonhyperactive	N	r
Age	40	.15
IQ	40	.01
Matching Familiar Figures - Latency	40	.46**
Matching Familiar Figures - Errors	40	-.36*
Self Control Rating Scale (Kendall, 1979)	40	-.16
Child Behavior Checklist Hyperactivity Factor (Achenbach, 1970)	40	.33*
Behavior Rating Scale	40	-.58***

Outpatient Clinic Sample-Nonhyperactive Subjects	N	r
Behavior Rating Scale (Conners, 1969) Hyperactivity Factor:		
Completed by Teachers	30	-.53**
Completed by Parents	30	-.50**
Child Behavior Checklist Hyperactivity Factor (Achenbach, 1970)	27	-.44**
Bender-Gestalt Impulsivity Score	29	.42**
Rorschach Human Movement (M)	31	.68**

*$p < .05$
**$p < .01$
***$p < .001$

To establish norms for the GDS Delay Task scores, a standardization study was recently completed in the public schools. The sample included a total of 192 boys and girls in three age groups (6-7 yrs.; 8-9 yrs.; 10-11 yrs.). It was determined that the mean Efficiency Ratio for the Delay Task (Total Rewards/Total Responses) was .84 (SD = .13) and that the scores were unaffected by age, sex, IQ and socioeconomic status. As with previous studies, performance remained consistent throughout the 8-minute session. To establish test-retest reliability, half the group were retested after 30 - 45 days, and it was found that the scores remained stable over that period. Norms are also available for the Vigilance Task of the GDS for a sample of 92 boys and girls.

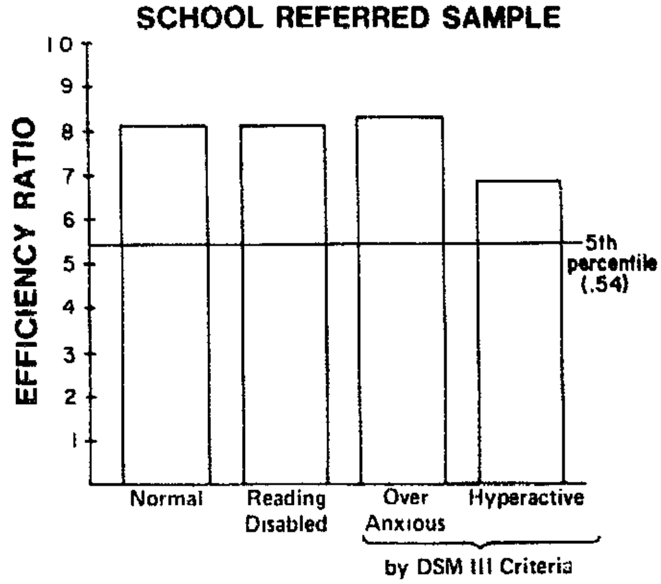
331

MEANS OF GDS VARIABLES: STANDARDIZATION SAMPLE

	6 7 year olds				8 9 year olds				10 11 year olds			
	Boys		Girls		Boys		Girls		Boys		Girls	
[illegible]ay Task (n = 192)	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Total Rewards	40 0	9 0	36 6	12 0	42 8	10 5	42 2	10 4	47 3	10 5	45 9	8 4
Total Responses	57 6	15 2	47 1	17 7	53 2	14 9	52 0	14 2	54 4	13 7	53 8	9 8
Total Efficiency Ratio	80	16	83	18	83	13	83	13	88	09	86	11
Vigilance Task (n = 91)	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Total Correct	30 7	9 4	25 7	11 8	37 6	6 0	38 5	3 7	41 8	1 7	41 5	2 3
Total Omissions	14 3	9 4	19 3	11 8	7 4	6 0	6 5	3 7	3 2	1 7	3 5	2 3
Total Commissions	16 1	15 5	9 3	5 0	11 8	10 7	9 6	10 3	6 9	8 0	5 1	4 1

Also completed recently was a study of 60 children who were classified by the school district as falling into one of four groups. Attention Deficit Disorder, reading disabled, overanxious and normal The GDS Delay and Vigilance Tasks successfully differentiated between the ADD group and the other three categories. It was also found that the correlations among the Delay Task Efficiency Ratio and the Vigilance Task scores fell in the .6 to .7 range of Pearson's correlation coefficient

MEAN EFFICIENCY RATIO: SCHOOL REFERRED SAMPLE



The hyperactive group had a lower mean Efficiency Ratio than the other 3 groups

While the GDS has successfully differentiated between groups classified as hyperactive and nonhyperactive, the accuracy of the results is often difficult

332

to assess because of the absence of reliable, independent criteria for making the classification of hyperactivity. Even when multiple criteria are employed, the degree of diagnostic agreement among the criterion measures tends to be poor. Consequently, when the GDS classifies a child differently, the other criteria may well be at fault. Indeed, when we have pursued such cases with follow-up interviews of parents and teachers, the GDS classification has, with few exceptions, seemed most plausible. (In one case, a girl classified as impulsive by the GDS had been classified nonhyperactive by the teacher's rating. When asked about the child, the teacher responded, "Oh, she's real hyper but she's so darling that I don't really think she's hyperactive.") Our experience to date has led us to orient future research away from criterion-based studies and towards those which focus on behavioral correlates of GDS performance.

The GDS will be further evaluated at 20 research sites in this country and Canada. In addition to expanding our data base, several projects will be investigating whether the device is sensitive to the effects of stimulant medication, while others will be examining the possibility of using it as a training procedure. There is also ongoing research into the possibility of including other tasks into the GDS which will enhance the yield of information.

REFERENCES

Achenbach, T. The child behavior profile: I. Boys aged 6-11. *Journal of Consulting and Clinical Psychology*, 1978, 46, 478-488.

Achenbach, T. and Edelbrock, C. The classification of children's psychopathology: A review and analysis of empirical efforts. *Psychological Bulletin*, 1978, 85, 1275-1301.

Cairns, E. and Cammock, T. Development of a more reliable version of the Matching Familiar Figures Test. *Developmental Psychology*, 1978, 14, 555-560.

Cohen, N., Douglas, V., & Morganstern, G. The effect of methylphenidate on attentive behavior and autonomic activity in hyperactive children. *Psychopharmacologia*, 1972, 22, 282-294.

Conners, C. A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry*, 1969, 126, 884-888.

Douglas, V. Stop, look, and listen: The problem of sustained attention and impulse control in hyperactive children. *Canadian Journal of Behavioral Sciences*, 1972, 4, 259-282.

Egeland, B. and Weinberg, R. The Matching Familiar Figures Test: A look at its psychometric credibility. *Child Development*, 1976, 47, 483-491.

Gardner, R. *The Objective Diagnosis of Minimal Brain Dysfunction*. Cresskill, N.J.: Creative Therapeutics, 1979.

Gordon, M. The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive boys. *Journal of Abnormal Psychology*, 1979, 7, 317-326.

Kagan, J. Reflectivity-impulsivity: The generality and dynamics of cognitive tempo. *Journal of Abnormal Psychology*, 1966, 71, 17-24.

Kendall, P. and Wilcox, L. Self-control in children: Development of a rating scale. *Journal of Consulting and Clinical Psychology*, 1979, 7, 317-326.

McClure, D.E. and Gordon, M. The performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. *Journal of Abnormal Child Psychology*, in press.

Machover, K. *Personality Projection in the Drawing of the Human Figure*. Springfield, Ill.: Charles C. Thomas, 1949.

Rosvold, H., Mirsky, A., Sarason, I., Bransome, E., & Beck. A continuous performance test of brain damage. *Journal of Consulting Psychology*, 1956, 20, 343-352.

333

Journal of Abnormal Child Psychology, Vol. 7, No. 3, 1979, pp. 317-326

The Assessment of Impulsivity and Mediating Behaviors in Hyperactive and Nonhyperactive Boys[1]

Michael Gordon[2]
The Ohio State University

Twenty boys (6-8 years) rated by their teachers as hyperactive and a matched sample of nonhyperactive boys performed a task that required them to withhold responding for a set time interval in order to be rewarded (DRL 6-second schedule). Half of each group worked on a one-button console while the other half was provided with additional collateral buttons. Results indicated that hyperactive children were relatively unable to perform efficiently on the task and that this deficit endured regardless of age, IQ, or experimental condition. DRL was thus found to discriminate accurately between teacher-rated and parent-rated hyperactive and nonhyperactive children. Furthermore, a wide variety of self-generated mediating behaviors was observed, and it was determined that a child's DRL performance was related to the kind of mediating behaviors he displayed. Results are discussed in terms of the clinical assessment of hyperactivity and the training of impulsive children.

Although the traditional approach has been to regard hyperactivity as a disorder of activity level, there is little evidence that the majority of children diagnosed as hyperactive are any more active, on the average, than a matched sample of nonhyperactive children, regardless of how activity is measured (Davis, Sprague, & Werry, 1969; Doubros & Daniels, 1966; Hutt, Jackson, & Level, 1966; Johnson, 1972; Lee & Hutt, 1964; McConnel, Cromwell, & Bialer, 1964; Ounsted, 1955; Schulman & Reisman, 1959). Moreover, viewing hyperkinesis solely in terms

Manuscript received in final form December 15, 1978.

[1] This report is based upon a dissertation submitted by the author in partial fulfillment for the Ph.D. degree. The efforts of Malcolm Helper, David Hothersall, Donald R. Meyer, Marilyn Moody, and Charles Wenar are gratefully acknowledged. The investigation was conducted while the author was a National Research Fellow, supported by Public Health Service Training Grant No. 5 T32 MH14608-02 to The Ohio State University.

[2] Address all correspondence to Michael Gordon, Upstate Medical Center, Division of Child and Adolescent Psychiatry, 750 East Adams Street, Syracuse, New York 13210.

0091-0627/79/0900-0317$03.00/0

of activity makes it difficult to understand many of the other symptoms that are often associated with this disorder, e.g., antisocial behavior, learning disabilities, distractibility, and excitability. The disenchantment with focusing on the quantity of a hyperactive child's activity has led some investigators to conceptualize hyperkinesis as a disturbance of impulse control and sustained attention. According to Douglas (1972), what typifies the hyperactive child is his inability to "stop, look and listen." She reviewed a body of research that appears to suggest that what characterizes the hyperactive child's behavior is not necessarily the sheer amount of activity, but its poor organization and control.

While the supporting studies cited by Douglas (1972) are intriguing, there is still need for a more straightforward assessment of the hyperactive's impulsivity. There is a tested technique, derived from research in operant conditioning, that provides an excellent opportunity to measure the hyperactive child's ability to inhibit behavioral responding. In this procedure, termed Differential Reinforcement of Low Rate Responding (DRL), responses that occur before a set time interval has passed are not reinforced and, moreover, they serve to reset the timer governing reinforcement. Thus a subject performing on a DRL 6-second schedule must wait at least 6 seconds between responses in order to receive a reinforcement. If a response is made before the 6 seconds have elapsed, the timer will reset and another 6 seconds must pass before a successful response can be made.

Normal laboratory animals performing on this task are able to develop efficient response patterns and typically come to be reinforced for 70% to 80% of the responses they make (Hothersall, Alexander, & Slonaker, 1972). As for human DRL performance, Weisberg and Tragakis (1967) demonstrated that even young children, ranging from 15 to 40 months, were able to maintain low levels of responding. Latency-age boys (Stein & Landis, 1973) and college students (Randolph, 1964) have also been shown to perform efficiently, although the samples for most studies were extremely small. One major question posed by the present study is: How would a hyperactive child fare in a DRL situation in which behavioral responding must be suppressed? Might the DRL procedure provide an objective measure of impulsivity?

Researchers investigating DRL performances have also observed that subjects in a time-based situation will often engage in a sequence or chain of behaviors between reinforceable responses. The notion is that this collateral behavior sequence helps the subject "wait out" the required temporal delay between responses. Examples of such behaviors in animals include wood chewing and tail licking (Slonaker & Hothersall, 1972). Similarly, human subjects have become convinced that these extra mediating behaviors may set up the reinforced responses (Bruner & Revusky, 1961). Experimentally precluding such collateral behaviors has led to a disruption of the subject's ability to make temporal discriminations (Stein & Landis, 1973, Stein & Flanagan, 1974). The

frequency of extra movements made by hyperactive children during a delayed reaction time task was positively related to the efficiency of their performances (Douglas, 1972; Cohen, cited in Douglas, 1972), suggesting that collateral behaviors may play a role in helping these children gain impulse control. Another intent of the current study was to examine whether or not a situation that facilitates the development of collateral behaviors helps to improve the DRL performance of hyperactive and, for that matter, normal children. And, regardless of the availability of experimentally introduced opportunities for collateral responding, will these children spontaneously elicit mediating responses?

METHOD

Subjects

Twenty hyperactive boys between 6 and 8 years of age (72-97 months) and a matched sample of nonhyperactive boys participated in the study. The subjects were selected from the client roster of a child guidance clinic; they were all nonretarded and enrolled in school at an age-appropriate grade level. A child was considered hyperactive if his teacher rated him as such on the hyperactivity factor of the Behavior Rating Scale (Conners, 1969). In accordance with the findings of a normative study (Werry, Sprague, & Cohen, 1975), those children with a total score of at least 15 on this scale were allotted to the hyperactive group, while subjects receiving a score of less than 15 made up the control group. The two groups were statistically comparable for age.

Design

The experiment essentially followed the design and methodology presented by Stein and Landis (1975). Four equal-sized groups of 10 subjects each were involved: two groups of hyperactive and two groups of nonhyperactive boys. One hyperactive and one nonhyperactive group performed on a DRL 6-second schedule without the availability of collateral buttons (DRL ONLY), while the other two groups worked on DRL using the five-button console (DRL + COLLATERAL)

Apparatus

Each subject was seated before a console containing either one (DRL ONLY) or five (DRL + COLLATERAL) microswitch buttons mounted 4.0 cm

apart. On a separate console, positioned immediately behind the button box, was a chassis containing a large six digit add counter which indicated the cumulative number of reinforcements (points) earned, and a red light, which lit for 2 seconds upon completion of a successful response. For the DRL ONLY condition, the one red button was programmed to produce reinforcement according to a DRL 6-second schedule. In the DRL + COLLATERAL condition, only one red button was programmed with the DRL schedule. The remaining four black collateral keys did not yield reinforcements, but they were individually monitored for the number of times they were pressed and the response pattern was retained by an Esterline Angus Event Recorder. All sessions were videotaped in their entirety. The programming equipment, counters, timers, recorders, and monitoring equipment were housed in a separate room, which was joined to the experimental room by a one way mirror and an intercom system.

Procedure

Prior to the testing session, the mother of every subject received an appropriate introduction to the study. The child was then taken individually to the experimental room, which contained only a table, chairs, and the testing apparatus. He then received the following instructions (adapted from Stein & Landis, 1975).

> You're going to play a game in which you will get a chance to win a lot of M&M's. Do you see this light [pointing to the reward indicator on the reinforcement box]? Every time you make this light go on you'll earn an M&M, and this counter [pointing] will keep track of how many M&M's you've won. At the end of the game I'll give you all the M&M's that you earned. Now, to make the light go on all you have to do is push this red button [pointing], and wait a little while before pressing it again. You just press this red button, wait a while, then press it again. If you press it again too soon, though, then you will not get a point. The light won't go on, and you'll have to wait a while before you can press it to get another point. But if you push the button, wait a while, then push it again, you'll get a point every time.

For the DRL + COLLATERAL groups, these instructions were added:

> These four black buttons will not give you points if you press them but, if you want, you can play with them while you're waiting to press the red button again.

The general instructions continued:

> Do you have any questions? [All questions were answered.] OK, I'll leave the room now and turn on the machine. When you hear my voice coming over the speaker saying, "You can start now," then you can begin playing the game. I'll come back when the game is over and bring you your M&M's. Have fun!

Each subject had a total of 15 minutes to earn points on the DRL schedule. At the session's end he was asked, "How did you do it?" and "Did you do

337

anything to help you wait?" Following his response, he was paid his due of M&M's, thanked for his participation, and returned to his mother.

Measures

The DRL task yields three measures of a subject's performance: the total number of responses, the number of reinforcements earned, and the efficiency score. Since this last score represents the percentage of reinforced responses, i.e., the number of reinforcements divided by the number of responses, it is considered to be the most sensitive indicator of a child's performance. Every subject's response, reinforcement, and efficiency scores were obtained for each of three 5-minute time blocks, and then totaled for the entire 15-minute session. Also monitored were the number of responses emitted on the collateral keys (if available) and the pattern in which they were pressed. Data on IQ, exact age, school grade, and parent ratings were also collected.

RESULTS

DRL Analysis

Analyses of variance (ANOVAs) were performed on the three operant measures, and the mean scores for each trial and for the whole session are shown in Figure 1. There was a highly significant main group (hyperactive vs. nonhyperactive) effect for all three DRL variables, especially the efficiency score, $F(1, 36) = 18.33$, $p < .0013$. Children rated as hyperactive by their teachers thus tended to emit more responses and receive fewer reinforcements during each of the time blocks than did children judged nonhyperactive. This main effect remains strong if the data are reanalyzed according to the parents' classifications of hyperactivity. The relative inefficiency of these hyperactive children suggests that they are significantly more impulsive as measured by the DRL 6-second schedule.

There were no Condition (DRL ONLY vs. DRL + COLLATERAL) effects on any of the dependent variables (efficiency score, $F(1, 36) = .133$, $p < .717$), so that whether a child played with the one-button or the five-button console had an insignificant effect upon his DRL performance. A few Group X Condition interactions did emerge. Hyperactive children in the DRL + COLLATERAL condition tended to make more responses than did their counterparts in the DRL ONLY condition, $F(1, 36) = 6.4$, $p < .016$, and nonhyperactive COLLATERAL subjects did significantly better according to the efficiency score than did the nonhyperactives in the DRL ONLY situation.

The ANOVAs yielded only one marginally significant main effect for Trials (time blocks). All groups tended to receive more rewards for their second-

338

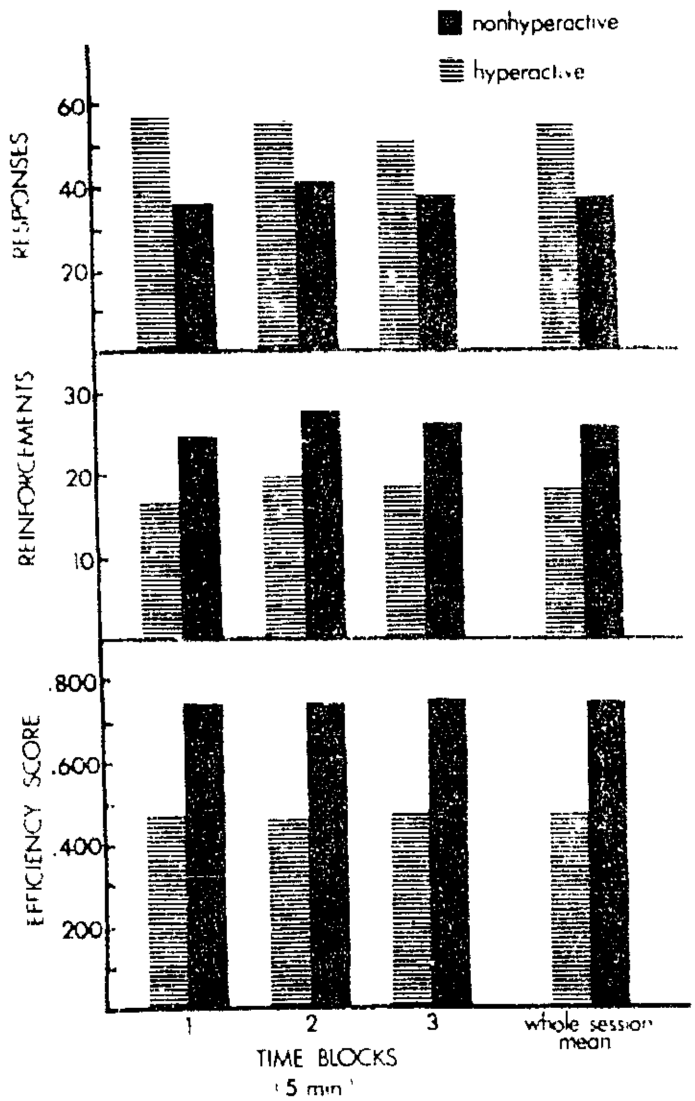


Fig. 1. The performance of hyperactive and nonhyperactive groups over [illegible] DRL [illegible] schedule. Top: responses; middle: reinforcements; bottom: efficiency [illegible]

339

trial performance, $F(2, 72) = 3.08$, $p < .051$. In general, the second trial was the most characteristic of a child's DRL work. For nearly all the various analyses, what held true for each trial taken separately, or for the session as a whole, usually held true for the second-trial measures and, again, particularly for the second time block efficiency score. This observation is supported by the intertrial correlations for each measure. Since the correlations between Trial 2 and Trial 3 tend to be quite high, mean $r = .69$, $p < .001$, it would suggest that a boy's performance was apt to stabilize by the second trial. When these interblock correlations are computed for each group separately, there is evidence that the DRL conduct of nonhyperactive subjects generally consolidated from one trial to the next as they locked into a relatively efficient pattern, mean $r = .75$, $p < .01$. The hyperactive children were more erratic across trials than the nonhyperactive children, mean $r = .54$, $p < .001$, although their performance was also relatively stable, as the intertrial correlations indicate.

In order to uncover relationships between the operant scores and a variety of descriptive data, all variables generated by this study were intercorrelated. It was first found that the subject's age is thoroughly unrelated to any other aspect of his performance or to the hyperactivity ratings supplied by the teacher or parent. The teacher ratings of hyperactivity using the Behavior Rating Scale (Conners, 1969) were significantly correlated with all DRL variables, their correlation with the whole session efficiency score, for example, was $r = .53$, $p < .01$. The parent rating was also tied to the major DRL scores (mean $r = .35$, $p < .05$), though slightly less than were the teacher ratings. The correlation between parents and teachers themselves, was $r = .61$, highly significant ($p < .001$), but not so high as to instill total confidence in their reliability. Parents and teachers did closely agree (34/40 or 85% of the cases) on the gross classification of a child as hyperactive or nonhyperactive. Of the six contesting cases, parents rated four teacher-classified nonhyperactives as hyperactive and two teacher-classified hyperactives as nonhyperactive. It seems, therefore, that parents were apt to classify more children as hyperactive. Yet the teachers' mean ratings of the hyperactive boys (22.4) was slightly higher than those rendered by parents (20.0). So, teachers classified fewer children as hyperactive, but they judged the hyperactives as more severely hyperactive.

Analysis of Mediating Behaviors

As has been the case for almost all animal and human subjects working on a DRL schedule, the children who participated in this study tended to engage in a variety of interresponse behaviors that were either physically manifested or were reported to the examiner in response to a standard, end-of-session question "How did you do it?" or "Did you do anything to help you wait?" A list of some of the mediating behaviors that were observed appears in Table

340

Table I. A List of Observed Mediating Behaviors

Circling DRL button with finger 9 times
Swinging legs 10, 12, or 20 times
Counting with lips
Counting out loud — numbers or ABC's
Blowing on reward box
Singing out loud
Shaking reward box 10 times
Hitting knee with right hand 20 times
Foot tapping 16 times
Tapping finger 10 times on button box
"Walking" fingers around DRL button 9 times
Stomping with foot 9 or 10 times
Running around table 1 time
Hitting side of box
Jumping jacks 4 times
Hitting collateral buttons

I. (It should be noted that only 5 of the 20 Collateral group used the collateral buttons for their interresponse regimens.) In order to be considered a mediating behavior, the sequence had to be repeated among at least 10 responses within a particular trial. Judgments concerning collateral responding were made blind from the videotapes. As it turned out, every subject engaged in, or reported, some mediating behavior. Ninety percent of the hyperactive subjects engaged in an observable mediating behavior, compared to 45% of the nonhyperactives. On the other hand, just 30% of the hyperactive boys reported the use of a nonbehavioral mediator, as opposed to 80% of the nonhyperactive subjects. A roughly comparable percentage of each group used both approaches (hyperactives = 20%, nonhyperactives = 25%).

There was an inverse correlation between the use of a physical collateral and both the efficiency ($r =$ .32, $p < .05$) and reinforcement ($r =$ .55, $p <$ .005) scores, especially after the first trial. Hence, children who used an observable collateral scheme did worse on the DRL schedule; they were also apt to be rated as more hyperactive by their parents and teachers. The opposite pattern emerged for cognitive mediators as there was a very high correlation between the children's reports of counting and all measures for each trial, particularly blocks 2 and 3, mean $r =$ .64, $p < .001$. Thus subjects who engaged in cognitive collaterals fared better on DRL and were usually judged less hyperactive than their noncounting counterparts. Boys who used both strategies during the session also had better scores, though not as high as the scores of those who only counted.

341

DISCUSSION

The prediction that children classified as hyperactive would encounter greater difficulties in managing the DRL task than would nonhyperactives was confirmed by the results of this study. According to every indicator of DRL performance, the hyperactive boys were relatively unable to refrain from making a high percentage of incorrect responses. Consequently, these results lend further credence to Douglas's (1972) proposal of an impulsivity dimension underlying the hyperkinetic syndrome, and to the validity of other measures of impulsivity presented by Sykes (cited in Douglas, 1972) and Campbell (cited in Douglas, 1972).

The experimental Condition (DRL ONLY vs. DRL + COLLATERAL) failed to emerge as a significant source of variance. However, given the observation that children engaged in many self-generated interresponse behaviors, this finding is not at all surprising. The fact is that a child will develop some type of interresponse strategy whether or not the experimenter provides a few extra buttons or telegraph keys. Indeed, in their attempts to quantify collateral responding, it appears that past investigators have missed the rich bounty of self-elicited behaviors that children exhibit during the course of a DRL session. Aside from their diversity, these schedule-induced behaviors are fascinating for the rigid manner in which they are executed; once a child chanced upon a successful strategy, he usually stuck with it ardently.

That the proclivity to engage in a physical or cognitive mediating behavior was so closely related to DRL performance as well as to ratings of hyperactivity is a highly intriguing result. Even with a relatively gross coding system, a strong pattern surfaced in which boys who displayed physical collaterals did worse on DRL and were rated more hyperactive than were boys who used cognitive means. In addition, many more hyperactive than nonhyperactive boys generated physical collaterals, an observation that may support the notion that a certain segment of the hyperactive child's hyperactivity is actually a series of mediating behaviors developed in an effort to control his impulses.

While it appears certain that cognitive mediators are more efficient strategies than physical ones, it is impossible to determine from these data whether physical collaterals actually hindered DRL efficiency or simply failed to facilitate it as much as the cognitive types did. The only hint available is the finding that the 18 subjects who employed physical collaterals during several trials had significantly higher efficiency scores than did the 9 children whose physical mediating was limited to one time block, $p < .05$. The precise nature of the interaction between a physical collateral and the DRL performance of hyperactive children is a very important avenue of research that warrants further exploration.

342

An unexpected aspect of these results was the extent to which the DRL variables could accurately discriminate between groups. According to the inter-trial correlations, the efficiency score can serve as a reliable and stable predictor of teacher-classified hyperactivity. Consequently, it is possible that this task could be developed into a useful clinical tool that could aid the diagnostician in the assessment of impulsivity and the classification of hyperactivity.

REFERENCES

Bruner, A., & Revusky, S. A. Collateral behavior in humans. *Journal of the Experimental Analysis of Behavior,* 1961, *4,* 349-350.

Cantwell, D. P. *The hyperactive child: Diagnosis, management, current research.* New York: Spectrum Publications, 1975.

Conners, C. K. A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry,* 1969, *126,* 152-156.

Davis, K., Sprague, R., & Werry, J. Stereotyped behavior and activity level in severe retardates: The effects of drugs. *American Journal of Mental Deficiency,* 1969, *73,* 721-727.

Doubros, S., & Daniels, G. An experimental approach to the reduction of overactive behavior. *Behaviour Research and Therapy,* 1966, *4,* 251-258.

Douglas, V. I. Stop, look and listen: The problem of sustained attention and impulse control in hyperactive and normal children. *Canadian Journal of Behavioral Science,* 1972, *4,* 259-282.

Gittis, A., & Hothersall, D. DRL performance of juvenile rats with septal lesions. *Physiological Psychology,* 1974, *2,* 38-42.

Hothersall, D., Alexander, D., & Slonaker, R. The DRL deficit of rats with septal lesions: Effects of extended training in a mediated environment. *Psychonomic Science,* 1972, *29*(1), 34-36.

Hutt, D., Jackson, P., & Level, M. Behavior parameters and drug effects: A study of a hyperkinetic epileptic child. *Epilepsia,* 1966, *7,* 250-259.

Johnson, C. F. Limits on the measurement of activity levels in children using ultrasound and photoelectric cells. *American Journal of Mental Deficiency,* 1972, *77,* 301-310.

Lee, D., & Hutt, C. A play-room designed for filming children: A note. *Journal of Child Psychology and Psychiatry,* 1964, *5,* 263-265.

McConnel, T. B., Cromwell, R. L., & Bialev, I. A. Studies in activity level: VIII. *American Journal of Mental Deficiency,* 1964, *68,* 647-651.

Ounsted, C. The hyperactive syndrome in epileptic children. *Lancet,* 1955, *269,* 303-311.

Randolph, J. J. A further examination of collateral behavior in humans. *Psychonomic Science,* 1964, *3,* 227-228.

Schulman, J. L., & Reisman, J. M. An objective measure of hyperactivity. *American Journal of Mental Deficiency,* 1959, *64,* 455-456.

Slonaker, R. L., & Hothersall, D. Collateral behaviors and the DRL deficit of rats with septal lesions. *Journal of Comparative and Physiological Psychology,* 1972, *80,* 91-96.

Stein, N., & Flanagan, S. Human DRL performance, collateral behavior and verbalization of the reinforcement contingency. *Bulletin of the Psychonomic Society,* 1974, *3,* 27-28.

Stein, N., & Landis, R. Mediating role of human collateral behavior during a spaced responding schedule of reinforcement. *Journal of Experimental Psychology,* 1973, *97,* 28-33.

Stein, N., & Landis, R. Differential reinforcement of low rates performance by impulsive and reflective children. *Journal of Experimental Child Psychology,* 1975, *19,* 37-50.

Weisberg, P., & Tragakis, C. J. Analyses of DRL behavior in young children. *Psychological Report,* 1967, *21,* 709-715.

Werry, J. S., Sprague, R. L., & Cohen, M. N. Conners' Teacher Rating Scale for use in drug studies with children: An empirical study. *Journal of Abnormal Child Psychology,* 1975, *3*(3), 217-229.

343

Journal of Abnormal Child Psychology, Vol. 12, No. 4, 1984, pp. 561-572

Performance of Disturbed Hyperactive and Nonhyperactive Children on an Objective Measure of Hyperactivity[1]

F. Daniel McClure
David C. Wilson Neuropsychiatric Hospital

Michael Gordon[2]
SUNY, Upstate Medical Center

Twenty hyperactive emotionally disturbed children (6-11 years) and a matched sample of nonhyperactive emotionally disturbed children were selected from the population of a therapeutic day treatment facility on the basis of teacher ratings. They were administered the Matching Familiar Figures Test-20 and were rated on several scales of impulsivity and/or hyperactivity. Each subject was required to perform on the Delay Task of the Gordon Diagnostic System, which required them to inhibit behavioral responding on a temporally based schedule (DRL-6) in order to win points. Children classified as hyperactive, whether by one or more criteria, were relatively unable to refrain from emitting a high number of nonreinforced responses. Moreover, these performance differences persisted regardless of age or IQ and were stable over the 8 minutes required to complete the test.

Manuscript received in final form January 4, 1984.
[1]The authors gratefully acknowledge the efforts of the following members of the Institute of Clinical Psychology, University of Virginia: Richard Abidin, Ann Loper, Ronald Reeve, and Sherry Kraft. Data analyses were conducted using STATMANAGER, Hayden Software Company, copyright 1983.
[2]Address all correspondence to Michael Gordon, Upstate Medical Center, 750 East Adams Street, Syracuse, New York 13210.

344

While current efforts to understand hyperactive behavior in children have highlighted deficits in attentional processes (Douglas & Peters, 1979), there has been mounting evidence that problems with impulse inhibition are also central to the disorder. In her review of the literature, Douglas (1972) concluded that hyperactive children "are apparently unable to keep their impulses under control in order to cope with situations in which care, concentration, or organized planning are required" (p. 275). Impulsive errors have been found to occur at substantially higher rates among hyperactive children on continuous performance tasks (Sykes, 1969; Sykes, Douglas, Weiss, & Minde, 1971; Sykes, Douglas, & Morganstern, 1973) as well as on delayed reaction time tasks (Cohen, 1970; Cohen, Douglas, Morganstern, 1972; Firestone & Douglas, 1975; Parry & Douglas, 1983), along with errors suggestive of lapses in attention. This dimension of impulsivity, which Douglas (1972) refers to as the "stop, look, and listen" dimension, is pervasive among these children and profoundly colors their behavior.

Measures of impulsivity in children currently available fall into two basic categories, those that are largely limited to laboratory investigations and those that were originally intended for the assessment of other behavioral domains but that have been adapted to assess impulsivity. Many clinicians rely upon adaptations of the Bender Gestalt test (Koppitz, 1964; Tolor & Branningan, 1980) and selected subtests of the WISC-R (Kaufman, 1979) to identify and assess impulsivity. The validity of these measures suffers because they assess a wide variety of rather loosely defined cognitive abilities that may or may not be affected by the dimension of impulsivity (Douglas, 1972).

Instruments such as the Continuous Performance Task (Rosvold, Mirsky, Sarason, Bransome, & Beck, 1956) and delayed reaction time tasks (Cohen, 1970; Cohen et al., 1972; Firestone & Douglas, 1975) have demonstrated a high degree of discriminative ability with hyperactive children, but have been largely limited to use in the laboratory. Rating scales such as the Teacher Rating Scale (Conners, 1969) have enjoyed widespread use in both research and clinical application with hyperactive children despite their inherent difficulties with rater bias, the "halo" effect, and the relative poor differentiation of the quantitative ratings involved (Ross & Ross, 1976).

Gordon (1979) has developed a methodology stemming from research in operant conditioning that has shown considerable promise as a stable and reliable measure of impulsive styles. This technique uses the Differential Reinforcement of Low Rate Responding (DRL) first described by Skinner (1938). The DRL schedule provides for reinforcement of a response that occurs after a certain specified time has elapsed since the emission of the previous response. Moreover, responses that occur prior to the termination of this time interval are not reinforced and serve to reset the timer governing the time interval. Thus, a subject performing on a DRL 6-second schedule would receive a reinforcement for every response emitted after an interval of 6 seconds, but responses emitted within 6 seconds of the previous response would not be reinforced, and the subject would have to wait another 6 seconds before the next response emitted could be reinforced.

In a study of 20 hyperactive boys and a matched sample of nonhyperactives, designated as such on the basis of the Teacher Rating Scale (Conners, 1969), Gordon (1979) found that this technique differentiated the two groups at a high level of statistical significance. The most sensitive measure was the efficiency ratio (ER), which is derived by dividing the total number of correct responses by the total number of responses. In addition, the nature of self-generated mediational strategies employed by each subject in his adaptation to the DRL schedule provided valuable data. Behaviors were rated as covert, or cognitive (e.g., counting silently), on the basis of posttest inquiry, or as overt, or behavioral, based on observer ratings. Statistical analysis of these collateral behaviors indicated that subjects who developed more cognitively oriented mediational strategies tended to achieve a significantly more efficient adaptation to the DRL schedule than did those subjects whose attempts at mediating were more behaviorally manifest.

An outgrowth of this research was the development of the Gordon Diagnostic System (GDS), which includes, in part, a version of the original DRL paradigm. (It also administers a version of the Continuous Performance Test.) In contrast to the electromechanical apparatus originally utilized, the GDS is a portable device that performs the DRL task and records the data in a single unit. This study was an attempt to determine if any performance differences would emerge in the use of this solid-state version of the task, and to extend Gordon's original study to include children with a wide variety of psychosocial, emotional, and psychological disturbances.

METHOD

Subject

Twenty emotionally disturbed children (18 males and 2 females) classified as hyperactive on the basis of the Teacher Rating Scale (TRS; Conners, 1969) and a matched sample of children rated as nonhyperactive (16 males and 4 females) were selected from the roster of a therapeutic day treatment facility.

346

Following from a normative study (Werry, Sprague, & Cohen, 1975), the criterion for designation as hyperactive was a mean rating of greater than 1.5 on the TRS. The two groups were statistically comparable for age, IQ, and SES (See Table I). While there was no method of controlling for severity of disturbances, the primary diagnoses for all children in this facility were conduct and emotional disorders, both in conjunction with cognitive and learning deficits.

Apparatus

In addition to the TRS, ratings were obtained from two other sources, the Self-Control Rating Scales (SCRS; Kendall & Wilcox, 1979) and the Teacher Report Form (TRF; Achenbach & Edelbrock 1983). The SCRS is a 33-item scale of impulsive behaviors, rated by teachers on a 7-point continuum. The TRF is a 113-item scale designed to assess a wide variety of behavioral problems and competencies. Behaviors are rated on a 3-point scale, with a rating of 0 representing a behavior that is "not true" of the child in contrast to a 2-point description of "very true."

Each child was administered the Matching Familiar Figures Test-20 (Cairns & Cammock, 1978), which is a longer, more reliable version of the original Matching Familiar Figures Test (Kagan, 1966). Each child also was administered the Bender Gestalt Test (Bender, 1938), which was scored for impulsivity (Tolor & Brannigan, 1980).

The Gordon Diagnostic System (GDS) is a one-button solid-state console manufactured by Clinical Diagnostics, Inc. It is designed to provide a behavioral measure of the child's ability to inhibit responding. The 6-second delay interval was employed in the current study. The GDS yields three fundamental sets of data: the absolute number of responses, the total number of rewarded responses, and an efficiency ratio (ER), which is obtained by dividing the total number of rewarded responses by the absolute number of responses. The ER thus represents the percentage of correct responses. If the child makes an unusally low or high number of responses, the ER may be a somewhat misleading index of performance. These three scores are computed for each 2-minute time block and then totaled for the overall 8-minute trial. During GDS performance each child's behavior was also recorded by the test administrator to assess the presence or persistence of any overt mediational strategies.

Procedure

Teachers initially completed the TRS and then, after a period of 4 weeks, the SCRS. The time interval was to minimize any possible overlap

347

Table I. Means and Standard Deviations for Hyperactive and Nonhyperactive Groups for Age, IQ, and TRS Factors

	Age in months	IQ WISC-R	Factor I	Factor II	Factor III	Factor IV
Hyperactive (*n* = 20)						
Mean	107.4	86.7	1.33	1.58	1.01	2.36
SD	26.2	11.85	.54	.60	.48	.42
Nonhyperactive (*n* = 20)						
Mean	102.9	82.9	.80	1.19	.93	1.19
SD	23.3	11.1	.48	.57	.43	.38

348

between ratings on the two scales. Head child care workers completed the TRF. The order of administration of the MFF-20 and the GDS was reversed for each group in order to control for the effects of boredom and/or fatigue. The Bender was always given first in the battery. The examiner was blind to each subject's group status.

Each subject was taken into the testing room individually and seated at a table. The MFF-20 was administered in standard fashion, and the GDS was introduced as a game, much like computerized arcade games. Prior to the GDS administration each subject was given the following instructions (adapted from Gordon 1979).

> Your are going to play a game in which you will get a chance to win a lot of points. Do you see this red light [pointing]? Every time you make this red light go on you'll earn a point and this counter [pointing] will keep track of how many points you've earned. Now, to make this light go on, all you have to do is push this blue button [pointing] and wait a while before pressing it again. You just push the blue button, wait a while, then press it again. If you press it too soon, though, you will not get a point, the red light won't go on, and you'll have to wait a while before you can press it to get another point. But, if you press the button, wait a while, and press it again, you'll get a point every time.

The examiner then answered any questions and made sure that the child had understood the instructions. Moving to a position adjacent to and somewhat behind the child the examiner continued:

> I'll sit here until you're done. You'll know when the game is over because this light will come on again. If you have any questions while you're playing the game please wait to ask them until the game is over. We can talk about the game when you're finished.

At the end of the trial each child was asked two questions: "Did you enjoy the game?" and "What did you do to help you wait?" The former was designed to address issues of motivation to perform well, while the latter was designed to address the issue of nonobservable mediational strategies.

RESULTS

To test the hypothesis that the GDS Delay Task could differentiate between the two groups, a one-tailed t test was conducted for each GDS measure and group membership (hyperactive vs. nonhyperative). The ER appeared to be the most sensitive measure, $t(38) = 5.67$, $p < .001$, and the absolute number of responses emitted also differentiated the two groups at a convincing level $t(38) = 4.26$, $p < .01$. Only the total number of rewarded responses failed to differentiate the two groups at a level of statistical significance, $t(38) = 1.79$, $p < .09$. While the two groups did not differ significantly on the number of times they were rewarded, those rated as hyperactive emitted a significantly greater number of nonrewarded

349

responses. The mean ER for the nonhyperactive group was .68 (SD = .15), while the mean ER for the hyperactive group was .42 (SD = .14). Since the ER is clearly derived from the other two scores, no comparisons of dependent variables were conducted.

To further judge the ability of the GDS to differentiate between children rated as hyperactive or nonhyperactive, post hoc analyses were performed with groups formed on the basis of being identified as hyperactive or nonhyperactive on four criterion measures (TRS, SCRS, MFF-20, and TRF) The Bender was not included as a criterion measure because the drawings of many of these subjects were so disturbed as to render them unscorable. The criterion for hyperactivity on the MFF-20 was a mean latency to first response of less than 10 seconds. Criteria for the TRF and SCRS were determined by taking the mean SCRS raw score and mean T score for the TRF hyperactivity factor for the entire sample, and then excluding all subjects within one standard deviation of the respective means. For subjects who were identified as hyperactive (n = 8) or nonhyperactive (n = 8) on the basis of all four criterion measures, there was agreement between the GDS and criterion measures in 15 cases (84%, Fisher's Exact Test, p = .01). The GDS criterion score was less than .54. When groups were constructed by subjects classified as hyperactive or nonhyperactive by *at least* three of the four criterion measures, there was agreement in 29 of 32 cases (91%, Fisher's Exact Test, p = .001).

Correlations presented in Table II depict the relationships between the three GDS variables and all the other measures. Although the criterion measures are related to one another to the extent of statistical significance or near statistical significance, only Factor IV of the TRS and both MFF-20 variables are significantly related to GDS performance.

Correlations between age and IQ for three GDS variables across each of the 2-minute time blocks as well as for the 8-minute trial yielded no statistically significant relationships. Correlations were also obtained for each GDS variable across each of the 2-minute blocks and for the total trial in order to assess the internal stability of the measure across time. While subjects rated as hyperactive demonstrated somewhat more variability in their performance across the 8 minutes, the patterns of responding were quite stable for both groups. Another method employed to look at the stability of the instrument in terms of discriminative ability was to run multiple t tests between the two groups for each GDS variable in each time block. Again, the absolute number of responses emitted and the ER consistently differentiated between the two groups, regardless of what particular discrete time block was used for comparison.

t tests conducted between TRS factors for both groups indicate that while Hyperactivity (Factor IV) was by far the most powerful in terms of differentiating the two groups, $t(38) = 9.28$, $p < .001$, the two groups also differed significantly along the dimensions of Conduct Problem (Factor I),

350

Table II. Correlational Matrices for the Three GDS Variables with All Other Measures

	Responses	Correct	FR	MFF-I	MFF-E	TRS I	TRS II	TRS III	TR-SIV	TRF	SCRS
Responses											
Total correct	31										
FR	82[b]	.62[a]									
MFF-I	.37[a]	.22	46[b]								
MFF-E	51[b]	− 11	67[a]	− 46[b]							
TRS I	16	−.20	28	01	25						
TRS-II	04	.08	02	− .09	12	19					
TRS III	35[a]	.16	.27	05	06	.22	47[b]				
TRS-IV	57[b]	−.21	59[b]	39[b]	23	43[b]	36	16			
TRF	17	−.05	26	.34	17	31	06	09	48[b]		
SCRS	02	.08	16	27	03	19	13	26	42[b]	56[b]	

[a] $p < .05$
[b] $p < .01$

$t(38) = 3.28$, $p > .001$, and Inattentiveness (Factor II), $t(38) = 2.11$, $p < .05$. This pattern of TRS scores is highly consistent with the pattern reported by Werry et al. (1975). Thus, it appears that although hyperactivity is the primary dimension measured by the TRS, it is also sensitive to related disturbances in attention and conduct. However, of these four factors, only Factor IV was significantly correlated with the GDS Efficiency Ratio, $r = .58$, $p < .01$.

Statistical analyses of mediational strategies were not conducted due to a paucity of data. Very few subjects engaged in any particular behavior(s) consistently enough for them to be considered mediational in nature, and few subjects were able to verbalize strategies that they may have employed.

DISCUSSION

The results lend considerable support to the GDS as a stable and reliable method of differentiating teacher-rated hyperactive from nonhyperactive children. Whether classified by the TRS alone or in conjunction with the other criteria, children classified as hyperactive were relatively unable to perform efficiently on this task. Although their average number of rewarded responses were not significantly different, the hyperactives were unable to refrain from emitting a high number of nonrewarded responses. Since there is a strong need for reliable behavioral measures in the overall assessment of hyperactivity (Prout & Ingram, 1982), the GDS may represent an important step forward.

Despite substantial qualitative differences in samples, the performances of the subjects in the present study very closely approximated those in Gordon's (1979) original study; both groups tended to obtain ERs lower than those reported earlier. Thus, it may be that both samples in this study actually consisted of hyperactive children who differed only in their degree of hyperactivity. However, if this is the case, the discriminative ability of the GDS is only enhanced, as it differentiated the two groups within a much more restricted range than would otherwise be the case. In the earlier study, the number of rewarded responses also differentiated the two groups at a statistically significant level. Here both groups also tended to make fewer rewarded responses. This difference may also relate to qualitative differences in samples between the two studies. Nevertheless, subjects rated as hyperactive in both studies demonstrated a significantly less efficient adaptation to the task than did their nonhyperactive counterparts.

The similarities in the performance of the subjects in this study and those in the original study persist despite some very important differences in subject selection and methodology. First, the miniaturization of the apparatus did not seem to affect the discriminative ability of the DRL procedure. Another important difference has to do with the nature of reinforcement. In the original study, each subject received an M & M for each rewarded response, and these were dispensed at the end of the session. In the present study, there were no tangible reinforcers.

Finally, this study was conducted with children whose emotional and behavioral difficulties are of such a magnitude as to require very restrictive education/therapeutic placement. Within this sample, the two groups differed significantly on three of the TRS factors—Conduct Disorder, Inattentive-Passive, and Hyperactivity. The relationships among these factors are consistent with the findings of several studies (Sprague et al., 1974) and would seem to support the notion that underlying deficits in attention and impulse inhibition contribute to behaviors that bring these youngsters into conflict with their environments The GDS appears to provide information about the extent to which a child, regardless of age, IQ, or specific diagnostic classification, displays such an "impulsive style."

The fact that, in this sample, age and IQ were unrelated in any way to Delay Task performance has very clear methodological and conceptual implications. A chronic difficulty with other measures of impulsivity, such as the Matching Familiar Figures Test (Cairns & Cammock, 1978), is that they are closely related to intellectual functioning and developmental status. Consequently, lack of consistency between measures of impulsivity and hyperactivity reported in some studies (Sandberg et al., 1978) may be related more to error of measurement than to the absence of a dimension of underlying impulsivity per se. It may also reflect that impulsivity is not a unitary construct (Paulsen & Johnson, 1980) and that the various tasks assess different aspects of a child's functioning as it relates to impulsive behavior. Children may be impulsive in different ways and to different extents as a result of variations in environmental, cognitive, and/or constitutional factors. In some children these deficits are expressed primarily through motoric channels while, in others, the expression may be more cognitive in nature.

It should also be pointed out that, although the Delay Task would seem to tap most broadly into the area of impulsivity, there are certainly a myriad of processes at play while the child performs the test. These most likely include motivational factors and time estimation ability, as well as the capacity to develop an efficient strategy. Further research into the interplay of these processes is in progress.

REFERENCES

Achenbach, T. M (1978). The child behavior profile I Boys aged 6-11 *Journal of Consulting and Clinical Psychology*, 46, 478-488

353

Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile.* New York. Queen City Printer.

Bender, L. (1938) A visual motor Gestalt test and its clinical. *American Orthopsychiatry Association Research Monographs, 3*

Carins, E., & Cammock, T. (1978). Development of a more reliable version of the Matching Familiar Figures Test. *Developmental Psychology, 14,* 555-560

Cohen, N. J. (1970). *Psychophysiological concomitants of attention in hyperactive children* Unpublished manuscript, McGill University.

Cohen, N. J., Douglas, V. I., Morganstern, G. (1972) The effect of methylphenidate on attentive behavior and autonomic activity in hyperactive children. *Psychopharmacologia, 22,* 282-294.

Conners, C. K. (1969). A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry, 126,* 884-888

Douglas, V. I. (1972) Stop, look, and listen. The problem of sustained attention and impulse control in hyperactive children. *Canadian Journal of Behavioral Sciences, 4,* 259-282

Douglas, V. I., & Peters, K. G. (1979) Toward a clearer definition of attentional deficits of hyperactive children. In G. A. Hale & M. Lewis (Eds.), *Attention and the development of cognitive skills* (pp. 173-247) New York, Plenum

Firestone, P., & Douglas, V. I. (1975) The effects of reward and punishment on reaction times and autonomic activity in hyperactive and normal children. *Journal of Abnormal Child Psychology, 3,* 201-215.

Gordon, M. (1979) The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive boys. *Journal of Abnormal Psychology, 7,* 317-326

Kagan, J. (1966) Reflectivity-impulsivity. The generality and dynamics of cognitive tempo. *Journal of Abnormal Psychology, 71,* 17-24.

Kaufman, A. S. (1979). *Intelligent Testing with the WISC-R.* New York. Wiley.

Kendall, P. C., & Wilcox, L. E. (1979) Self-control in children. Development of a rating scale. *Journal of Consulting and Clinical Psychology, 7,* 317-326

Koppitz, E. M. (1964) *The Bender Gestalt Test for Young Children.* New York. Grune and Stratton.

Parry, P. A., & Douglas, V. I. (1983). Effects of reinforcement on concept identification in hyperactive children. *Journal of Abnormal Child Psychology, 11,* 327-340.

Paulsen, K., & Johnson, M. (1980) Impulsivity: A multidimensional concept with developmental aspects. *Journal of Abnormal Child Psychology, 8,* 269-277

Prout, H., & Ingram, R. E. (1982). Guidelines for the behavioral assessment of hyperactivity. *Journal of Learning Disabilities, 15,* 393-395.

Ross, D. M., & Ross, S. A. (1976) *Hyperactivity: Research, theory, action.* New York: Wiley.

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, E., & Beck, L. A. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology, 20,* 343-352.

Sandberg, S. T., Rutter, M., & Taylor, E. (1978). Hyperkinetic disorder in psychiatric clinic attenders. *Developomental Medical and Child Neurology, 20,* 279-299

Skinner, B. F. (1938) *The behavior of organisms.* New York. Appleton-Century-Crofts

Sykes, D. H. (1969) *Sustained attention in hyperactive children.* Unpublished doctoral dissertation, McGill University

Sykes, D. H., Douglas, V. I., & Morganstern, G. (1973) Sustained attention in hyperactive chil-children. *Journal of Child Psychology and Psychiatry, 14,* 213-220

Sykes, D. H., Douglas, V. I., Weiss, G., & Minde, K. K. (1971) Attention in hyperative children and the effect of methylphenidate (Ritalin). *Journal of Child Psychology and Psychiatry, 12,* 129-139

Tolor, A., & Brannigan, G. (1980). *Applications of the Bender-Gestalt Test.* Springfield, Illinois: Thomas

Werry, J. S., Sprague, R. L., & Cohen, M. N. (1975) Conners' teacher rating scale for use in drug studies with children – An empirical study. *Journa. of Abnormal Child Psychology, 3,* 217-229

354

ASSESSMENT OF ATTENTION DEFICIT DISORDERS USING THE GORDON DIAGNOSTIC SYSTEM

Michael Gordon, Ph.D.
State University of New York, Upstate Medical Center

F. Daniel McClure, Ph.D.
David C. Wilson Neuropsychiatric Hospital

Presented at the 92nd Annual Convention of the American Psychological Association at Toronto, Ontario, Canada, August 1984

Correspondence address:
750 East Adams Street
Syracuse, New York 13210
(315) 473-8145

ASSESSMENT OF ATTENTION DEFICIT DISORDERS USING THE GORDON DIAGNOSTIC SYSTEM

Abstract

Despite large numbers of referrals for evaluations of Attention Deficit Disorders/Hyperactivity, school psychologists have had to contend with assessments conducted without the aid of objective and reliable procedures. The Gordon Diagnostic System (GDS) administers tasks which provide data on a child's ability to delay and to sustain attention (Gordon, 1979; Gordon & McClure, 1983; Gordon & McClure, 1984; McClure & Gordon, 1984). We have now gathered normative data on 700 children between 3 and 16 years of age. Studies are reported which compare diagnoses made by traditional approaches with those based on GDS performance. An analysis of the time expended in traditional assessment activities is also presented. In addition, we describe a study contrasting the GDS performances of children classified as severely learning disabled, educationally mentally handicapped, and emotionally handicapped. Finally, we introduce initial normative data on a new GDS task which provides information about the effects of distraction on sustained attention. The cost effectiveness of traditional evaluations as well as implications for future diagnostic evaluations are explored.

ASSESSMENT OF ATTENTION DEFICIT DISORDERS USING THE GORDON DIAGNOSTIC SYSTEM

INTRODUCTION

According to a recent survey (Gordon & McClure, 1983), approximately 30% of referrals to school psychologists for psychological testing contain a question of hyperactivity. Of these referrals, practioners reported that only about 17% were eventually diagnosed as "truly" hyperactive. It was documented that the school psychologists relied far more on clinical judgment than they did on data from psychological tests which they considered relatively irrelevant to the diagnosis. Comments by the respondents reflected a general lack of confidence in the evaluation of this disorder.

The Gordon Diagnostic System (GDS) was introduced to provide practitioners with a means of obtaining accurate data on a child's ability to delay and to sustain attention. It has been shown that the GDS accurately discriminates between groups classified as hyperactive and nonhyperactive in samples from an outpatient clinic (Gordon, 1979), a day treatment center for severely emotionally disturbed children (McClure & Gordon, 1984), and a school-referred population (Gordon & McClure, in press). It also has been shown that the GDS Delay Task is sensitive to the effects of stimulant medication (Shue & Douglas, 1983).

DESCRIPTION OF THE GDS

The GDS is a portable, electronic unit which allows for the administration of two tasks. The Delay Task requires the child to inhibit responding in order to gain a point. Specifically, the child is instructed to press a button, wait, and then press the button again. If he/she refrains from responding for the delay interval, a light flashes and a reward counter increments. If the child responds before the interval elapses then the timer resets and no reward points are recorded. For children 6 years of age and older, a delay interval of 6 seconds is used and the session lasts 8 minutes. Preschoolers are required to wait 4 seconds across a 6-minute session.

The GDS Delay Task generates three raw scores: the number of responses (button presses), the number of correct responses (i.e. Rewards) and the Efficiency Ratio which represents the percentage of correct responses.

The GDS also contains the Vigilance Task which is a version of the Continuous Performance Test (Rosvold, et al., 1956). The child is presented with a series of digits on the front display and is told to press the button every time a "9" appears that had been preceded by a "1". The GDS records the number of correct presses, the number of times the child fails to press the button when the "1/9" combination appears (i.e. errors of omission), and the number of extraneous button presses (i.e. errors of commission). For the testing of younger children, the GDS contains a "1" mode, which requires the subject to press the button every time a "1" appears. The same performance measures are recorded. The Vigilance Task lasts 9 minutes for children 6 years and older, and 6 minutes for preschoolers.

The GDS Distractibility Task is a new procedure we are developing to assess the effects of distraction on a child's ability to sustain attention. It is identical to the Vigilance Task except that random numbers appear at random intervals on each side of the "hot" digit. The same set of scores are recorded for this task as for the Vigilance Task.

A microcomputer software package is also available that automatically receives data from the GDS, tabulates the various scores, stores the information, and prints out an interpretive summary of the performance. Included in the report are educational recommendations and suggestions for pharmacotherapy.

THE GDS IN THE ELEMENTARY SCHOOL -- Current Studies

Prior research (Gordon, 1979; Gordon & McClure; 1983,1984; McClure & Gordon, 1984) established preliminary normative data and demonstrated the diagnostic accuracy of the GDS. We now have greatly expanded the

[redacted]

standardization sample and have gathered data on clinical use of the GDS in actual school settings.

[redacted]

A. The Standardization Study -- Normative data for the Delay and Vigilance Tasks are based upon the testing of 700 children while pilot data for the Distractibility Task come from the performances of 58 children. Subjects were selected at random from nine public and private schools in New York State and Virginia. To be included in the sample the child had to be rated by the teacher as nonhyperactive on the Teacher Rating Form (Edelbrock & Achenbach, 1984) or Teacher Rating Scale (Conners, 1969), and have no history of grade retention. Data on IQ, achievement, and SES were also collected.

For ease of interpretation, we have arranged the standardization data into Threshold Tables (see Tables 1-3). Based upon data from standardization projects, we have divided scores into three ranges. A score is considered ABNORMAL if it is typical of less than 5 percent of the normal population, i.e. the 5th percentile or less. Scores in the BORDERLINE range are those which fall between the 6th and 25th percentile. Finally, scores that are above the 25th percentile are classified as NORMAL. These cutoffs correspond closely to clinical experience and our research studies in that children classified as hyperactive by other measures typically perform in the ABNORMAL range of GDS scores.

B. The Cost-Effectiveness and Accuracy of Traditional Evaluations for ADD/Hyperacitvity -- Prior studies have shown that the GDS classifies children as hyperactive or nonhyperactive with a high level of accuracy when behavior rating scales and research tasks are used as criteria (McCLure & Gordon, 1984). At the same time, numerous studies have demonstrated that traditional assessment techniques such as the WISC-R and the Bender-Gestalt Designs, do a poor job at discriminating between groups classified as hyperactive and nonhyperactive (Sandoval, 1977). The pitfalls of relying on clinical judgment have also been documented (Hughes, Goldman, & Snyder, 1983; Marquis, 1983).

To explore the ramifications of current practices in the evaluation of ADD/Hyperactivity, we asked 3 highly-experienced school psychologists to record the amount of time they spent conducting assessments of 31 children referred with hyperactivity as a presenting complaint. The instruments administered during the course of their assessments included the Wechsler Intelligence Scale for Children-Revised (WISC-R), Bender Gestalt Designs, Children's Personality Questionnaire, Peabody Individual Achievement Test, Visual Aural Digit Span Test, Sentence Completion, House-Tree-Person, and Draw-A-Person. The results, presented in Table 4, indicate that each evaluation occupied between 10 and 16 hours of the psychologist's time.

Although their time was distributed across several categories of activity, they spent the most number of hours actually testing the child.

After the psychologists had arrived at a diagnosis based on their assessment, each child was tested with the GDS. The children were classified as ADD by the GDS criteria if their scores fell within the ABNORMAL range on one or both of the GDS Tasks. The agreement between classifications conducted with and without the GDS is illustrated in Table 5. Using the GDS scores as criteria, the results indicate that evaluations by traditional assessment tended to seriously overdiagnose ADD. The full assessment battery did not appear to miss children who were truly ADD but that finding may have been affected by the nature of the sample. Upon further inquiry after completion of the study, it appeared that the children considered by the GDS scores to be misdiagnosed actually represented those cases about which the clinicians were least confident in their classifications.

C. The GDS Performance of Children in Self-Contained Classrooms -- The purpose of this study was to determine the prevalence of attentional disorders in a sample of children placed in self-contained classrooms. The data presented here actually represent the initial findings of a large study of GDS performance in school-based populations of retarded, blind, deaf and emotionally disturbed children. The children in the current sample had been labelled by the

Committee on the Handicapped as either Severely Learning Disabled (SLD; n=13), Educationally Mentally Handicapped (EMH; n=7), or Emotionally Handicapped (EH; n=9). The results (Table 6) indicated that nearly a quarter of the sample scored abnormally on both Tasks and would be considered significantly impulsive and inattentive. The highest percentage of abnormal GDS performances came from the EH group where nearly 55% had one or both scores in the ABNORMAL range. Inspection of teacher rating scales and discussions with the teachers themselves indicated that approximately 75% of this group had overactivity as a presenting complaint. The teachers also suggested that information about a particular child's level of impulsivity and inattentiveness helped them develop more systematic remediation programs.

Summary

Our early investigations with the GDS centered mostly around clinic populations, and the preliminary norms were based on a relatively small sample. The normative data are now quite extensive and we have had considerable experience with using the GDS in school settings. The findings presented here indicate that school psychologists, using traditional assessment techniques, expend a great deal of time evaluating children for ADD/Hyperactivity. There is also the suggestion that classifications based on traditional assessment lead to an overdiagnosis of the problem. This would seem to indicate the possibility that many children classified as ADD/hyperactive either did not require treatment or might have benefitted from other forms of remediation. Finally, we have found that children in self-contained classrooms can be evaluated with the GDS. A significant percentage of the sample showed signs of impulsivity and inattentiveness.

ACKNOWLEDGMENTS

These studies were possible because of the cooperation of the Liverpool School District, Liverpool, NY. Thanks are due particularly to Mr. Kevin Bacher, Mr. William Mugridge, Ms. Susan Venditto, Ms. Gerry Mouio, and Mr. William Cleary.

We are greatly indebted to Ms. Barbara Mettelman, Research Coordinator for GDS Research Projects, Dr. Ernest M. Post, Dr. Wendy Gordon, Mr. Jeffrey Weiner, Ms. Sandy Tavitian, and Ms. Lacie DeHimer for their active involvement in our research program.

This research was funded by a grant to the Research Foundation of the State University of New York awarded by Clinical Diagnostics, Inc.

NOTE:

The GDS is commercially available from Clinical Diagnostics, Inc., 1536 Cole Blvd. Suite 350, Golden, Colorado 80401 (Phone: 800-521-4503).

References

Conners, C.K. (1969). A teacher rating scale for use in drug studies with children. American Journal of Psychiatry, 126, 884-888.

Edelbrock,C. & Achenbach, T.A. (1984). The Teacher Version of the Child Behavior Profile: I. Boys aged 6-11. Journal of Consulting and Clinical Psychology, 52, (2), 207-217.

Gordon, M. (1979). The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive boys. Journal of Abnormal Child Psychology, 7 (3), 317-326.

Gordon, M. & McClure, F.D. (1983, August). The objective assessment of attention deficit disorders. Paper presented at the 91st Annual Convention of the American Psychological Association, Anaheim, California.

Gordon, M. & McClure, F.D. (1983). Survey of current practices and problems in the assessment of attention deficit disorders. Manuscript submitted for publication.

Hughes, M.C., Goldman, B.L., Snyder, N.F. (1983). Hyperactivity and the attention deficit disorder. American Family Physician, 27, 119-126.

Marquis, P. (1983). The attention deficit disorder. Postgraduate Medicine, 73, 295-300.

McClure, F.D. & Gordon, M. (1984). Performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. Journal of Abnormal Child Psychology, 12, (4), 561-571.

Rosvold, H.E., Mirsky, A.F., Saranson, I., Bransome, E. & Beck, L.A. (1956). A continuous performance of brain damage. Journal of Clinical Psychology, 20, 343-352.

Sandoval, J. (1977). The measurement of the hyperactive syndrome in children. Review of Educational Research, 47, 293-318.

Shue, K. & Douglas, V.I. (1983). [The effect of Ritalin on hyperactive children's performance on a DRL measure of impulsivity]. Unpublished raw data.

Table 1. Threshold Table for the Delay Task

3-5 YEAR OLDS (n=138)

	TOTAL ER	TOTAL REWARDS	TOTAL RESPONSES
NORMAL	.61 - 1.0	≥ 46	≤ 64
BORDERLINE	.39 - .60	16 - 25	65 - 91
ABNORMAL	≤.38	≤ 15	≥ 92 & ≤ 20

[redacted]

6-7 YEAR OLDS (n=139)

	TOTAL ER	TOTAL REWARDS	TOTAL RESPONSES
NORMAL	.73 - 1.0	≥ 35	≤ 62
BORDERLINE	.42 - .72	14 - 34	63 - 75
ABNORMAL	≤.41	≤ 13	≥ 76 & ≤ 20

8-11 YEAR OLDS (n=269)

	TOTAL ER	TOTAL REWARDS	TOTAL RESPONSES
NORMAL	.79 - 1.0	≥ 43	≤ 66
BORDERLINE	.55 - .78	31 - 42	67 - 76
ABNORMAL	≤.54	≤ 30	≥ 77 & ≤ 20

12-16 YEAR OLDS (n=154)

	TOTAL ER	TOTAL REWARDS	TOTAL RESPONSES
NORMAL	.86 - 1.0	≥ 47	≤ 66
BORDERLINE	.71 - .85	39 - 46	67 - 75
ABNORMAL	≤.70	≤ 38	≥ 76 & ≤ 20

Table 2. Threshold Tables for the Vigilance Task

[illegible]-5 YEAR OLDS (n=138)

	TOTAL CORRECT	TOTAL COMMISSIONS
NORMAL	21 - 29	0 - 5
BORDERLINE	13 - 20	6 - 24
ABNORMAL	≤ 12	≥ 25

6-7 YEAR OLDS (n=137)

	TOTAL CORRECT	TOTAL COMMISSIONS
NORMAL	33 - 45	≤ 10
BORDERLINE	22 - 32	11 - 23
ABNORMAL	≤ 21	≥ 24

8-9 YEAR OLDS (n=136)

	TOTAL CORRECT	TOTAL COMMISSIONS
NORMAL	40 - 45	≤ 8
BORDERLINE	33 - 39	9 - 21
ABNORMAL	≤ 32	≥ 22

10-11 YEAR OLDS (n=133)

	TOTAL CORRECT	TOTAL COMMISSIONS
NORMAL	42 - 45	≤ 3
BORDERLINE	38 - 41	4 - 15
ABNORMAL	≤ 39	≥ 16

12-16 YEAR OLDS (n=154)

	TOTAL CORRECT	TOTAL COMMISSIONS
NORMAL	43 - 45	≤ 3
BORDERLINE	40 - 42	4 - 11
ABNORMAL	≤ 39	≥ 12

Table 3. Threshold Table for the Distractibility Task.

7-14 YEAR OLDS (N=58)

	TOTAL CORRECT	TOTAL COMMISSIONS
NORMAL	28 - 45	≤ 6
BORDERLINE	6 - 27	7 - 35
ABNORMAL	≤ 5	≥ 36

Table 4. Time Analysis of Traditional Evaluations for ADD/Hyperactivity.

Number of cases evaluated	31
Mean number of hours spent:	
Administering Tests	3 - 5
Conducting Classroom Observations & Teacher Meetings	2 - 3
Interviewing Parents	1 - 2
Scoring Tests	1 - 2
Writing Reports	3 - 4
Total time expended (in hours)	10 - 16

Table 5. Accuracy of traditional evaluations using GDS scores as criteria.

	ADD	NON-ADD
HITS	13	6
MISSES	0 (FALSE NEGATIVES)	12 (FALSE POSITIVES)

Table 6. Percentage of children with ADD/Hyperactivity in self-contained classrooms

(N=29)

	n	ABNORMAL DELAY	ABNORMAL VIGILANCE	BOTH TASKS ABNORMAL
SLD	13	2 (15 %)	5 (38 %)	2 (15 %)
EMH	7	2 (29 %)	3 (43 %)	1 (14 %)
EH	9	5 (56 %)	6 (67 %)	4 (44 %)
TOTAL	29	9 (31 %)	14 (48 %)	7 (24 %)

(Psychopharmacology Bulletin)

Microprocessor-Based Assessment of Attention Deficit Disorders

Dissatisfaction with formulating a diagnosis of ADD/Hyperactivity based almost entirely upon clinical judgment or the perception of raters has spawned the development of behavior-based assessment procedures. These efforts typically involve administration of the Continuous Performance Test (Rosvold, Mirsky, Sarason, Brasome,Jr., & Beck, 1956), or related measures of attention and self-control (Davenport, 1972; Doyle, Anderson & Halcomb, 1976; Hiscock, Kinsbourne, Caplan & Swanson, 1979; Margolis, 1972). Although bulky and expensive electromechanical devices had previously been required, researchers are now programming microcomputers to administer these laboratory tasks to children (Conners, 1980; Klee & Garfinkel, 1983).

Use of the microcomputer has certain advantages, including flexibility of administration, the ability to store multiple data points, and, if a microcomputer is already available, cost-effectiveness. This approach, however, does have drawbacks for the researcher and, in particular, the clinician. The size of even portable microcomputers can make testing in multiple sites cumbersome. It is perhaps for this reason that standardization samples for software-driven programs tend to be, at best, limited in number and breadth. The transport of disk drives, monitors, keyboards, and the computer, itself, from location to location often discourages

371

use of the procedure in other than a single research or clinic setting.

Another potential limitation of microcomputer-based testing stems from concerns surrounding the reliability of administration. Computer monitors vary in the size, shape, intensity and color of displayed characters. Each brand of computer presents to the child a different keyboard, and accepts different types of peripherals, such as "joysticks, which are often used as manipulanda. Without control over such critical factors as the amount of pressure necessary to actuate a switch, or the accuracy of timed sequences, repeatable administration across multiple microcomputers cannot be assured.

Finally, ADD/Hyperactive children, despite a reputation for academic underachievement, routinely display a fine facility for disassembling delicate equipment. Unless the examiner is immediately present, the highly impulsive child will inevitably stick a finger in disk drives, unplug cables, adjust monitors, or in some other fashion interrupt standard administration of the task. Efforts to secure the microcomputer are often unsatisfactory because they usually involve additional hardware, cabling, and expense.

An alternative approach to testing attention and self-control has been developed by the author (Gordon & McClure, 1983; Gordon & McClure, 1984). The goal of this project has been to establish a practical, reliable, and well-standardized procedure available to both researchers and clinicians. The Gordon Diagnostic System (GDS) is a microprocessor-based, portable unit which, without an external microcomputer, allows for the administration of multiple tasks. The Delay Task, based upon a DRL operant schedule requires the child to inhibit responding in order to gain a reward. Specifically, the child is instructed to press a button, wait, and then press the button again. If s/he refrains from responding for at least six seconds, a light flashes and a reward counter increments. If the child responds before the interval elapses then the timer resets and no reward points are recorded. The GDS Delay Task generates three major scores: the number of responses (button presses), the number of correct responses (i.e. Correct) and the Efficiency Ratio which represents the percentage of correct responses.

The GDS also contains the Vigilance Task, which is a version of the Continuous Performance Test (Rosvold, et al., 1956). The child is presented with a series of digits on the front display and is told to press the button every time a "9" appears that had been preceded by a "1". The GDS records the number of correct presses, the number of times the child fails to press the button when the "1/9" combination appears (i.e. errors of omission), and the number of extraneous button presses (i.e. errors of commission). For the testing of younger children, the GDS contains a "1" mode, which requires the subject to press the button every time a "1" appears. The same performance measures are recorded. X

Although normative data were gathered using standard settings for task parameters, the design allows the user to select a wide range of parameters. This feature enabled the modification of parameters for the testing of adults as well as very young children.

Because the GDS is controlled by a programmable microprocessor, additional tasks can be explored. One such enhancement being studied is a version of the Vigilance Task which assesses the impact of distraction on a child's ability to sustain attention. This Distractibility Task is identical to the standard Vigilance Task except that random digits flash at random intervals on the outer positions of the LED display. The subject is still required to press the blue button when a "9" comes right after a "1". The only difference is that numbers flash on either side of the center (i.e. relevant) digit. Other enhancements under development are designed to evaluate attention across sensory modalities.

Even though the GDS administers tasks independently of a microcomputer, it can communicate with external hardware via an RS232C communications port. Software is available which allows for the direct transmission of GDS data to a microcomputer where it can be tabulated, graphed, stored, and compared to normative data. Connection of the GDS to a microcomputer enables collection of some ancillary data which cannot be extracted from the stand-alone unit.

The portability and ease-of-operation of the GDS has allowed for large-scale data collection. The standardization sample is comprised of 900 boys and girls from 3 to 16 years of age (Gordon & Mettelman, 1985). An additional 1100 hyperactive and nonhyperactive protocols from various subject populations, including the deaf, blind, Spanish-speaking, emotionally disturbed, and learning disabled, have also been gathered.

A series of validational studies has shown that these game-like tasks differentiated accurately between hyperactive and nonhyperactive children from both outpatient, and day treatment settings (Gordon, 1979; McClure & Gordon, 1984). In a sample of school-referred children, the GDS distinguished children with ADD from those classified as reading disabled, overanxious, and normal (Gordon & McCLure, 1983). Over 20 university and medical center research sites are currently conducting investigations involving the GDS. Most studies concern the effectiveness of the GDS for determining drug responsivity and for evaluating the success of pharmacotherapy (Atkinson, 1985; Barkley, 1985; Rapport, 1985). Others are investigating the use of the GDS for evaluating children with known brain damage, preschoolers considered "at-risk" for impulsive behavior, and the relationship between GDS scores and observational measures of classroom behavior. The GDS is also being used by school systems and pediatric practices across the country for the clinical evaluation of ADD/Hyperactivity.

COMMENTS AND CONCLUSIONS

A chronic impediment in the field of ADD/Hyperactivity has been the paucity of universal criteria and procedures for subject selection and treatment monitoring. Even in the realm of behavioral assessment, where truly creative approaches to testing children have been developed, procedures have generally not been sufficiently practical or well-standardized to allow for acceptance by the majority of professionals (Loney, 1981). As such, most continue to rely on rating scales or tests, such as the WISC-R (Wechsler, 1974), which have been shown to be of limited usefulness in differentiating between groups (Douglas, 1972). Furthermore, comparison of studies across research programs has often been hampered by idiosyncratic measures or sets of criteria.

The GDS represents an effort to establish a standard procedure for evaluating certain aspects of attention and self-control. While it is more costly and, in certain respects, less flexible than software-driven approaches, its portability, ruggedness, and extensive base of normative data offer the user greater assurance of reliable administration and valid interpretation. The growing acceptance of the GDS within the professional community also permits greater comparison of data across studies.

It must be emphasized, however, that the GDS was never intended as a divining rod for ADD/Hyperactivity. This is a complex disorder which represents an array of subgroups and interactive diagnostic dimensions. The GDS is seen as an important tool to be used only in conjuction with other selection criteria and clinical judgment.

References

Atkinson, A.W., Cohen, P.C. & Kelly, P.C. (1985, April). Attention deficit disorder: The effects of ritalin on self-esteem a comparison of ACTeRS teacher scale, Conners' parent scale, and Gordon Diagnostic System in diagnosis and management. Paper presented at the American Academy of Pediatrics Meeting, Atlanta, Georgia.

Barkley, R.A. (1985,August). Assessment of stimulant drug responding in ADD-H children. Paper presented at the 93rd Annual Convention of the American Psychological Association, Los Angeles, California.

Conners, C.K. (1980). Continuous performance test [Computer program]. Kensington, Maryland: Behavioral Medicine, Inc.

Davenport, W. (1972). Vigilance and arousal: Effects of different types of background stimulation. The Journal of Psychology, 82, 339-346.

Douglas, V.I. (1972). Stop, look and listen: The problem of sustained attention and impulse control in hyperactive and normal children. Canadian Journal of Behavioral Science, 4, 259-282.

Doyle, R.B., Anderson, R.P. & Halcomb, C.G. (1976). Attention deficits and the effects of visual distraction. Journal of Learning Disabilities, 9, 59-65.

Gordon, M. (1979). The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive children. Journal of Abnormal Psychology, 7, 317-326.

Gordon, M. & McClure, F.D. (1983). The objective assessment of attention deficit disorder. Paper presented at the 91st Annual Convention of the American Psychological Association, Anaheim, California.

Gordon, M. & McClure, F.D. (1984). Assessment of attention deficit disorders using the Gordon Diagnostic System. Paper presented at the 92nd Annual Convention of the American Psychological Association, Toronto, Canada.

Gordon, M. & Mettelman, B.B. (1985). Threshold tables for the Gordon Diagnostic System. (Available from Clinical Diagnostics, Inc., 300 E. Mineral Avenue, Suite 6, Littleton, Colorado 80122).

Hiscock, M., Kinsbourne, M., Caplan, B. & Swanson, J.M. (1979). Auditory attention in hyperactive children: Effect of stimulant medicaiton on dichotic listening performance. Journal of Abnormal Psychology, 88, 27-32.

Loney, J. (1981). Evaluating treatments for childhood hyperactivity: Some methodological considerations. In K.D. Gadow & J. Loney (Eds.) Psychosocial aspects of drug treatment for hyperactivity (pp. 77-103). Boulder, Colorado: Westview Press.

Klee, S.H. & Garfinkel, B.D. (1983). The computerized continuous performance task: A new measure of inattention. Journal of Abnormal Child Psychology, 11(4), 489-496.

Margolis, J.S. (1972). Academic correlates of sustained attention. Unpublished thesis, University of California, Los Angeles.

McClure, F.D. & Gordon, M. (1984). The performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. Journal of Abnormal Child Psychology, 12(4), 561-572.

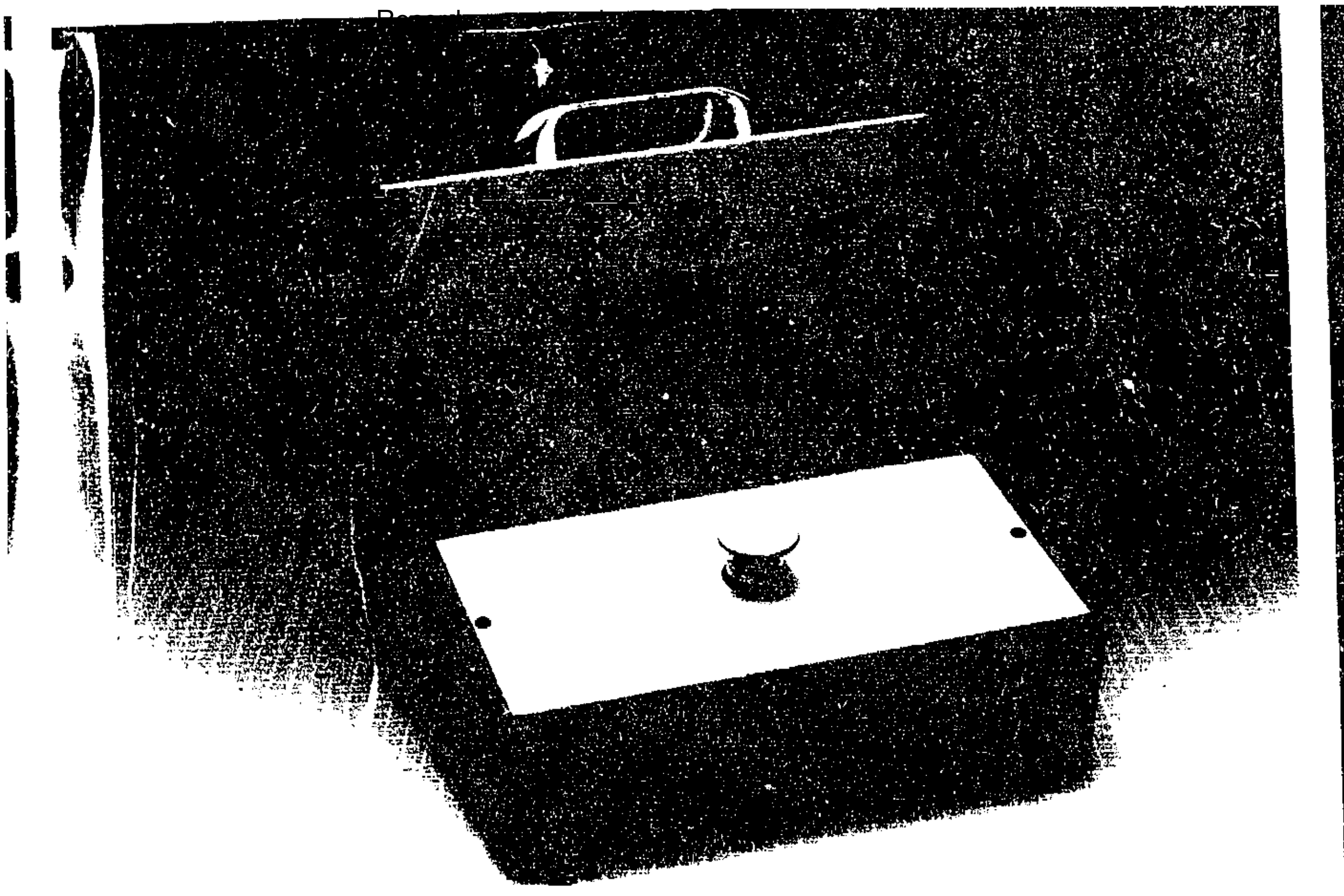
Rapport, M.D. (1985, August). Comparing classroom and clinic measures of ADD: Dose response effects. Paper presented at the 93rd Annual Convention of the American Psychological

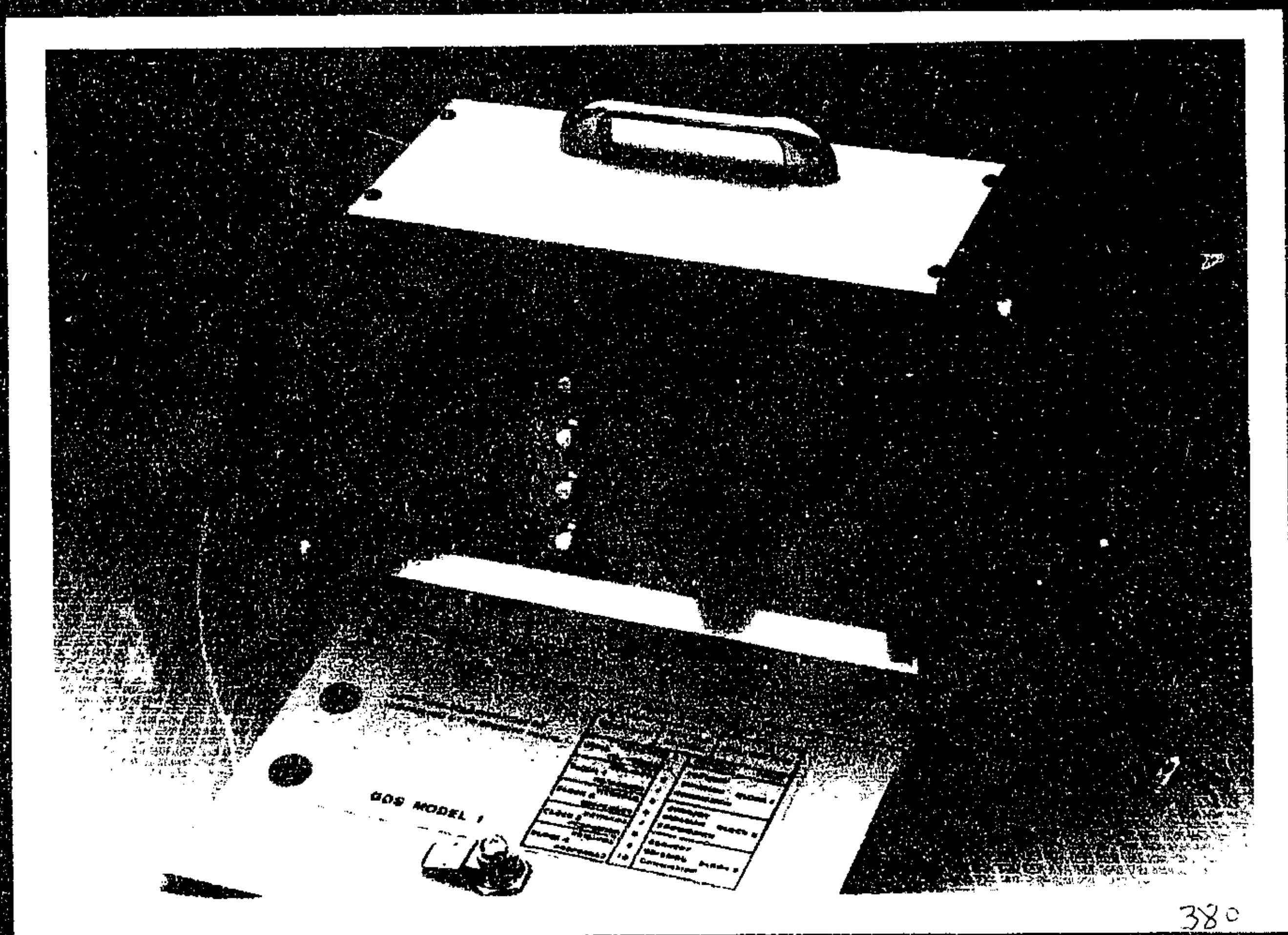
Association, Los Angeles, California.

Rosvold, H.E., Mirsky, A.F., Sarason, I., Bransome,Jr., E.D. & Beck, L.H. (1956). A continuous performance of brain damage. Journal of Consulting Psychology, 20, 343-350.

Weschsler, D. (1974). Weschsler intelligence scale for children-revised. New York: Psychological Corporation.

379





380

510(k) Review

K854903 B

Company Name CLINICAL DIAGNOSTICS, INC.

Device Name GORDON DIAGNOSTIC SYSTEM MODEL 1

	YES	NO
1. Life-supporting or life-sustaining?		✓
2. Implant (short-term or long-term)?		✓

3. Similar preenactment device(s): CONTINUOUS PERFORMANCE TEST (ROSVOLD) and other psychological performance tests
(device name, manufacturer)

4. Differs from preenactment devices how?

(1) Psychological performance tests were not considered by any FDA classification panel and no classification regulation or proposal has been written.

(2) It is not clear whether these devices are medical devices or not.

(3) Labeling and intended use appear to meet the definition of a "device," therefore, I believe the 510(k) is appropriate.

5. If appropriate: provides comparative in vitro data: see publication
provides a summary of animal testing? N/A
provides a summary of clinical testing? extensive

6. I believe this is equivalent to device(s): # none
Classification should be based on:

~~Subsection~~ UNCLASSIFIED (name) (presently Class ——).

(sign & date) [signature] 5/86

I believe this is not equivalent to any preenactment device. ____________

I believe clinical testing is required before a determination can be made. ____________

 RFM Chron

MEMORANDUM OF TELEPHONE CONVERSATION

Between: Millie Clark
Clinical Diagnostics
(303) 795-0438

And: Chief, Neurological Devices Branch, HFZ-430
Office of Device Evaluation, DANRD

Date: June 2, 1986

Subject: K854903 B

Summary: I was told Bruce Meyer was out of town. I was referred to Millie Clark, who said she was acting on behalf of Mr. Meyer regarding this 510(k).

I pointed out that the labeling might not meet the requirements of § 801.109. She agreed to do the following:

1. a precise statement of intended use would be provided under "indications." This statement would describe the device as an aid in diagnoses, and would not imply that it could determine a diagonoses without other information.
2. References to specific drug regimens would be deleted from the device labeling.

Robert P. Munzner, Ph.D.

cc:

4

DEPARTMENT OF HEALTH & HUMAN SERVICES — Public Health Service

Memorandum

Date: May 29, 1986

From: REVIEWER(S) - NAME(S) R.F. Munzner Ph.D.

Subject: 510(k) NOTIFICATION K854903/B

To: THE RECORD

It is my recommendation that the subject 510(k) Notification:

✓ (A) Is substantially equivalent to marketed devices.

____ (B) Requires premarket approval. NOT substantially equivalent to marketed devices.

____ (C) Requires more data.

____ (D) Is an incomplete submission. (See Submission Sheet).

Additional Comments:

The submitter requests:

✓ No Confidentiality

____ Confidentiality for 90 days

____ Continued Confidentiality exceeding 90 days

Class Code w/Panel: 84LQD

ATTENTION TASK PERFORMANCE RECORDER

REVIEW: [signature] Munzner, PhD (BRANCH CHIEF) 5/30/86 (DATE)

FINAL REVIEW: [signature] (DIVISION DIRECTOR) 5/30/86 (DATE)

DO NOT REMOVE THIS ROUTE SLIP!!!!

K-85-4903

4/30/86

FROM: CLINICAL DIAGNOSTICS, INC. ATTN: BRUCE MEYER 300 E. MINERAL AVENUE, #6 LITTLETON, CO 80122	LETTER DATE 11/25/85 TYPE OF DOCUMENT: 510 (k)	LOGIN DATE 12/09/85	DUE DATE 05/21/86 CONTROL # K854903
TO: ODE/DMC	CONT. CONF.: ? STATUS : H REV PANEL : ~~HO~~ NE PAN/PROD CODE(S): ~~HO~~/NE / /		
SUBJECT: GORDON DIAGNOSTIC SYSTEM GDS DATA ANALYSIS PROGRAM			
DECISION: DECISION DATE: / /	RQST INFO DATE: 02/10/86 DATE: 04/30/86 DATE: / /	INFO DUE DATE: 03/12/86 DATE: 05/30/86 DATE: / /	

SUPPLEMENT: 01 LTR DATE: 860219 LOGIN DATE: 860220

DEPARTMENT OF HEALTH & HUMAN SERVICES — Public Health Service

Memorandum

Date 28 April 1986

From REVIEWER(S) - NAME(S) Mattan / Hinckley

Subject 510(k) NOTIFICATION K854903/A

To THE RECORD

It is my recommendation that the subject 510(k) Notification:

_____ (A) Is substantially equivalent to marketed devices.

_____ (B) Requires premarket approval. NOT substantially equivalent to marketed devices.

X (C) Requires more data.

_____ (D) Is an incomplete submission. (See Submission Sheet).

Additional Comments:
Hold for additional information.
See enclosed phone memo.

The submitter requests:

X No Confidentiality

_____ Confidentiality for 90 days

_____ Continued Confidentiality exceeding 90 days

Class Code w/Panel: _____

REVIEW: ________________
(BRANCH CHIEF) (DATE)

FINAL REVIEW: ________________
(DIVISION DIRECTOR) (DATE)

DEPARTMENT OF HEALTH & HUMAN SERVICES Public Health Service

Memorandum

Date MAR 10 1986

From REVIEWER(S) - NAME(S) Mattan

Subject 510(k) NOTIFICATION K854903/A

To THE RECORD

It is my recommendation that the subject 510(k) Notification:

✓ (A) Is substantially equivalent to marketed devices.

___ (B) Requires premarket approval. NOT substantially equivalent to marketed devices.

___ (C) Requires more data.

___ (D) Is an incomplete submission. (See Submission Sheet).

Additional Comments:

MEDICAL COMPUTERS AND SOFTWARE

See Memo

Use old letter

The submitter requests:

✓ No Confidentiality

___ Confidentiality for 90 days

___ Continued Confidentiality exceeding 90 days

Class Code w/Panel: 80 LNX unclassified

REVIEW: [signature] (BRANCH CHIEF) 3/14/86 (DATE)

FINAL REVIEW: [signature] (DIVISION DIRECTOR) 3/14/86 (DATE)

MEMORANDUM OF TELEPHONE CONVERSATION

Between: Mr. Bruce G. Meyer Jr.
President
Clinical Diagnostics, Inc.
300 E. Mineral Ave. #6
Littleton, Colorado 80122
(303) 795-0438

And: Physiologist, DANRD, HFZ-430

Date: 28 April 1986

Subject: Premarket Notification K854903A

Mr. Meyer and I discussed the information he submitted in the above noted submission concerning the Gordon Diagnostic System (GDS). During our conversation he said he would forward the following information to HFZ-401 labeled as an addendum to this file:

(b)(4)

[redacted]

(b)(4)

[redacted]

2

230

MEMO RECORD	AVOID ERRORS PUT IT IN WRITING	DATE MAR 12 1986
FROM: Biomedical Engineer		OFFICE ODE
TO: Record		DIVISION DGGD
SUBJECT: Clinical Diagnostics Inc. GDS and DAP		

SUMMARY

The Gordon Diagnostic System (GDS) is a portable electronic device designed to assess deficits in attention and impulse control in children. It has been developed for use by clinicians as an aid in the diagnosis of attention deficits disorders ~~with~~ as well as some forms of learning disabilities. The GDS administers two tasks: 1) The Delay Task measures a child's ability to refrain from pressing in order to win points and; 2) the Vigilance Task assesses how well a child sustains attention over a long period of time.

The GDS Data Analysis Program (DAP) is an optional software program designed to simplify the collection of data generated by the GDS, by organizing and storing the data.

A child is placed in one of three categories (normal, borderline, abnormal) based on the comparison of his scores to the norms established through a standardization study.

The two tasks performed in the GDS are based on psychological testing instruments developed for research. The Continuous Performance Test ('56) is the basis for the GDS Vigilance Task, and the Differential Reinforcement of Low Rate Respon-

SIGNATURE	DOCUMENT NO. K854903

FORM FDA 2034 (12/80)

(b)(6)
ding ^ is the basis for the GDS Delay Task.

Amalie C. Mattan

DO NOT REMOVE THIS ROUTE SLIP!!!!

K-85-4903

2/10/86

FROM: CLINICAL DIAGNOSTICS, INC. ATTN: BRUCE MEYER 300 E. MINERAL AVENUE, #6 LITTLETON, CO 80122	LETTER DATE 11/25/85	LOGIN DATE 12/09/85	DUE DATE 03/09/86
	TYPE OF DOCUMENT: 510 (k)		CONTROL # K854903
TO: ODE/DMC	CONT. CONF.: ? STATUS : H REV PANEL : HO PAN/PROD CODE(S): HO/ / /		
SUBJECT: GORDON DIAGNOSTIC SYSTEM GDS DATA ANALYSIS PROGRAM			
DECISION: DECISION DATE: / /	RQST INFO DATE: 02/10/86 DATE: / / DATE: / /	INFO DUE DATE: 03/12/86 DATE: / / DATE: / /	

DEPARTMENT OF HEALTH & HUMAN SERVICES Public Health Service

Memorandum

Date FEB - 1986

From REVIEWER(S) - NAME(S) Mitton

Subject 510(k) NOTIFICATION K854903

To THE RECORD

It is my recommendation that the subject 510(k) Notification:

____ (A) Is substantially equivalent to marketed devices.

____ (B) Requires premarket approval. NOT substantially equivalent to marketed devices.

✓ (C) Requires more data.

____ (D) Is an incomplete submission. (See Submission Sheet).

Additional Comments:

Hold

The submitter requests:

✓ No Confidentiality

____ Confidentiality for 90 days

____ Continued Confidentiality exceeding 90 days

unclassified
Class Code w/Panel: 80 LNX

REVIEW: ________________ (BRANCH CHIEF) (DATE)

FINAL REVIEW: ________________ (DIVISION DIRECTOR) (DATE)

MEMO RECORD	AVOID ERRORS PUT IT IN WRITING	DATE FEB 7 1986 10 40 AM
FROM: Biomedical Engineer		OFFICE ODE
TO: Record		DIVISION DGGD

SUBJECT: K854903 Gordon Diagnostic System GDS DAP

SUMMARY

I spoke with Mr. Bruce Meyer, Jr., President, to request the following additional information:

① The similarities and differences between the GDS and the products to which Clinical Diagnostics is claiming S.E (Continuous Performance test, CPT, and the Differential Reinforcement of Low Rate Responding, DRL);

② The basis for the educational recommendations; and

③ a description of and a basis for the "suggestions for pharmacotherapy" (Mr Meyer said that this will not be a feature of the GDS DAP, and will include a statement to that effect with ~~to~~ the additional information)

Mr. Meyer will send the additional info to DMC and is aware that the file will be on hold until that info is received.

SIGNATURE	DOCUMENT NO.
Amalie C. Mattson	K854903

FORM FDA 2034 (12/80)

DEPARTMENT OF HEALTH AND HUMAN SERVICES **Public Health Service**

Food and Drug Administration
Center for Devices and
Radiological Health
8757 Georgia Avenue
Silver Spring, MD 20910

FEBRUARY 20, 1986

CLINICAL DIAGNOSTICS, INC.
ATTN: BRUCE MEYER
300 E. MINERAL AVENUE, #6
LITTLETON, CO 80122

D.C. Number : K854903
Received : 02-20-86
Product : GORDON DIAGNOSTIC SYSTEM GDS DATA ANALYSIS PROGRAM

The additional information you have submitted has been received.

We will notify you when the processing of your submission has been completed or if any additional information is required. You are required to wait ninety (90) days after the received date shown above or until receipt of a "substantially equivalent" letter before placing the product into commercial distribution. I suggest that you contact us if you have not been notified in writing at the end of this ninety (90) day period before you begin marketing you device. Written questions concerning the status of your submission should be sent to:

Food and Drug Administration
Center for Devices and
Radiological Health
Office of Device Evaluation
Document Mail Center (HFZ-401)
8757 Georgia Avenue
Silver Spring, Maryland 20910

If you have procedural or policy questions, please contact the Division of Small Manufacturers Assistance at their toll-free number (800) 638-2041 or me at (301) 427-8162

Sincerely yours,

Robert I. Chissler
Premarket Notification Coordinator
Office of Device Evaluation
Center for Devices and
Radiological Health

- NEW ADDRESS -
300 E. Mineral Avenue, #6
Littleton, Colorado 80122
303-795-0438

K854903/A

Clinical Diagnostics, Inc.

BASIS MEDICAL PRODUCTS

February 19, 1986

Mr. Robert Gatling
General Hospital and Personal
Use Devices
Food and Drug Administration
Document Mail Center (HFZ-401)
8757 Georgia Avenue
Silver Spring, Maryland 29010

Dear Mr. Gatling:

Attached is our response to the requests for additional information by telephone from Amalie. Please expedite action on our 510K application as quickly as possible. Clinical Diagnostics has been marketing the Gordon Diagnostics System for the past two years, and based on directives from your department we are now holding all orders. I would greatly appreciate a telephone call as to the current status.

Thank you very much for your kind cooperation.

Sincerely,

CLINICAL DIAGNOSTICS, INC.

Bruce G. Meyer, Jr.
President

BGM:em

1536 COLE BOULEVARD, SUITE 350, GOLDEN, CO 80401
(303) 232-7129
(800) 521-4503

Clinical Diagnostics, Inc.

Ref: K854903
Product: Gordon Diagnostic System
GDS Data Anaylsis Program

In response to the request for additional information:

(b)(4)

[redacted]

1536 COLE BOULEVARD, SUITE 350, GOLDEN, CO 80401

(303) 232-7129
(800) 521-4503

(b)(4)

[redacted]

Sincerely,

CLINICAL DIAGNOSTICS, INC.

Bruce G. Meyer, Jr.
President

BGM:em

DEPARTMENT OF HEALTH AND HUMAN SERVICES **Public Health Service**

Food and Drug Administration
Center for Devices and
Radiological Health
8757 Georgia Avenue
Silver Spring, MD 20910

MAY 1, 1986

CLINICAL DIAGNOSTICS, INC.
ATTN: BRUCE MEYER
300 E. MINERAL AVENUE, #6
LITTLETON, CO 80122

D.C. Number : K854903
Received : 05-01-86
Product : GORDON DIAGNOSTIC SYSTEM GDS DATA ANALYSIS PROGRAM

The additional information you have submitted has been received.

We will notify you when the processing of your submission has been completed or if any additional information is required. You are required to wait ninety (90) days after the received date shown above or until receipt of a "substantially equivalent" letter before placing the product into commercial distribution. I suggest that you contact us if you have not been notified in writing at the end of this ninety (90) day period before you begin marketing you device. Written questions concerning the status of your submission should be sent to:

Food and Drug Administration
Center for Devices and
Radiological Health
Office of Device Evaluation
Document Mail Center (HFZ-401)
8757 Georgia Avenue
Silver Spring, Maryland 20910

If you have procedural or policy questions, please contact the Division of Small Manufacturers Assistance at their toll-free number (800) 638-2041 or me at (301) 427-8162

Sincerely yours,

Robert I. Chissler
Premarket Notification Coordinator
Office of Device Evaluation
Center for Devices and
Radiological Health

STATE UNIVERSITY OF NEW YORK
UPSTATE MEDICAL CENTER
750 EAST ADAMS STREET
SYRACUSE, NEW YORK 13210

K854903/B

AREA CODE 315
473-8100

COLLEGE OF MEDICINE
DEPARTMENT OF PSYCHIATRY AND BEHAVIORAL SCIENCES

April 30, 1986

Food and Drug Administration
Center for Devices and Radiological Health
Office of Device Evaluation
Document Mail Center
(HFZ - 401)
8757 Georgia Avenue
Silver Spring, Maryland 20910

DC#: K854903
Product: Gordon Diagnostic System

To whom it may concern:

Enclosed please find materials requested by Mr. Stephen Hinckley regarding our submission for FDA approval.

I would appreciate a rapid forwarding of these documents to Mr. Hinckley.

Sincerely,

Michael Gordon

Michael Gordon, Ph.D.
Associate Professor of Psychiatry

STATE UNIVERSITY OF NEW YORK

UPSTATE MEDICAL CENTER
750 EAST ADAMS STREET
SYRACUSE, NEW YORK 13210

COLLEGE OF MEDICINE
DEPARTMENT OF PSYCHIATRY AND BEHAVIORAL SCIENCES

AREA CODE 315
473-8100

April 30, 1986

Mr. Stephen Hinckley
Food and Drug Administration
8757 Georgia Avenue
Silver Spring, MD 20910

Dear Mr. Hinckley:

Enclosed please find a series of reprints regarding the various issues we discussed today by phone. You will note that I have highlighted sections that would seem most relevant to the equivalency issue. Also included are instructions —manuals and advertisements to similar programs that are marketed for clinical use. Perhaps the most current and useful information for your purposes is found in the short reviews from Psychopharmacology Bulletin.

The articles by Rapport and Barkely, cited in an earlier response to the FDA review, are also included. Finally, there is a basic description of the standardization process and a copy of a page from a widely—used psychological test, which demonstrates the conventionality of our statistical approach to presenting normative data.

I hope that, in the rush to send these materials to you, we have not forgotten anything essential. I have tried to be somewhat conservative in selecting articles so that that the approval process does not become further delayed by a bad case of eye strain.

I enjoyed talking with you and trust you will contact me with any further questions or comments.

Sincerely,

Michael Gordon

Michael Gordon, Ph.D.
Associate Professor of Psychiatry

7

Attention Vigilance Performance

attention Response Task recorder

Attention Task Response Reaction

ARTICLES RELEVANT
TO EQUIVALENCY
ISSUE

ORIGINAL CPT ARTICLE

Journal of Consulting Psychology
Vol. 20, No. 5, 1956

A Continuous Performance Test of Brain Damage[1,2]

H. Enger Rosvold,[3] Allan F. Mirsky,[3] Irwin Sarason,[4]
Edwin D. Bransome, Jr.,[5] and Lloyd H. Beck[6]
Yale University

This paper presents the results of an investigation using a new instrument for the study of brain damage in human subjects. The design of the instrument, the Continuous Performance Test (CPT) was based on certain electroencephalographic evidence which suggested that brain-damaged individuals should show inferior ability as compared with non-brain-damaged individuals on tasks requiring sustained attention or alertness.

The waking EEGs of brain-damaged patients generally show either random bursts of hypersynchronous (high amplitude) activity intruding upon the normal activity of the brain, or a general hypersynchrony (3, 8). Hypersynchronous activity is also evident in the recording from the brain of a sleeping subject (3). If hypersynchrony is associated with reduced vigilance or attention, as suggested by its presence during sleep, then the hypersynchrony of the brain-damaged patients might also indicate reduced attention. For example, according to Hebb (1), hypersynchrony might interfere with the sequential firing of cell assemblies and thereby disrupt the process of attention. Other research strongly suggests that some such relationship does in fact exist. Thus, on the basis of a study of 75 prefrontal lobotomy cases, Levin states:

> Immediately following operation the patient usually lapsed into a drowsy-anergic state from which he could be temporarily aroused by stimuli, only to lapse once again into somnolence when the stimuli were removed. During this drowsy-anergic state EEGs showed diffuse slow activity and a good deal of baseline oscillations of ½ to 2 per second anteriorly mixed with slow rhythms of 2 to 6 per second frequency. As the patient emerged from the drowsy stage the preoperative pattern usually asserted itself (2, p. 422).

The usual measures of attention or alertness such as the digit span and digit symbol substitution subtests of the Wechsler-Bellevue have not consistently showed decline following brain damage (4, 5, 6, 7). In the case of the patient who shows intermittent bursts of hypersynchronous activity, and hence only momentary lapses in attention, the test performance may be due to the fact that the *S* can to a great extent choose his own time to respond, and may reorganize his attention between momentary lapses. The lapses would then not affect his score to any measurable extent.

In the case of the individual who shows a generalized hypersynchrony of the electrical

[1] The authors would like to extend their appreciation to the following individuals and institutions for providing advice and/or subjects and testing facilities for this research: Mr. Samuel Greenhouse of the National Institute of Mental Health, Bethesda, Maryland; Dr. C. Edward Stull and the Southbury Training School, Southbury, Connecticut; Drs. Walter Landmesser and Russell Fuldner and the Newington Crippled Children's Hospital, Newington, Connecticut; Dr. Rhoades and the New Haven Cerebral Palsy Clinic, New Haven, Connecticut; Drs. Milton Senn, William German, Janice Stevens, Morris Wessell, Ethelyn Klatskin, and the staff of the several pediatric and neurological clinics of the Grace-New Haven Community Hospital, New Haven, Connecticut.

[2] Supported in part by grants to H. Enger Rosvold from the Veteran's Administration, Contract VA1001-M3222 and the National Science Foundation.

[3] Now at the National Institute of Mental Health, Bethesda, Maryland.

[4] Now at the West Haven V.A. Hospital, West Haven, Connecticut.

[5] Now at the College of Physicians and Surgeons, Columbia University, New York, N. Y.

[6] Now at the University of Michigan, Ann Arbor, Michigan.

activity of the brain, the classical tests might not reflect his deficiency because they do not require sustained attention over a sufficiently long period of time.

These considerations suggested that a test which would not allow the patient to choose his own time to respond, and which would require a high level of continuous attention over an appreciable interval of time might reflect a deficit that other procedures would miss. Accordingly, the CPT was designed to provide two attention tasks, labeled *X* and *AX*, with *AX* designed to be the more difficult. This paper presents data on the CPT performance of comparable groups of brain-damaged and non-brain-damaged individuals.

If the research demonstrates that there are indeed differences between brain-damaged and non-brain-damaged groups on this test, one of the questions which arises is whether or not the differences are large enough to be of diagnostic use. The data also provide a preliminary answer to this question.

Method

Subjects. No attempt was made in this study to select patients on the basis of either locus, extent, or type of brain damage. The term "brain-damaged" is here used to include any individual for whom there was medical evidence (surgical notes, neurological examination, EEG) of brain pathology. Three groups were studied, each consisting of a brain-damaged and a non-brain-damaged subgroup. The *Defective* group consisted of 72 institutionalized feebleminded *Ss*, who were diagnosed as either of "organic etiology" (the *Brain-Damaged-Defective* subgroup) or of "familial or idiopathic etiology" (the *Non-Brain-Damaged-Defective* subgroup). The *Child* group comprised 45 children from a children's hospital and from various pediatric clinics who were being treated for brain disorders such as cerebral palsy (the *Brain-Damaged-Child* subgroup) or for non-nervous-system disorders (the *Non-Brain-Damaged-Child* subgroup). The third group, designated the Adult group, consisted of 50 adult *Ss*. Half of these (the *Brain-Damaged-Adult* subgroup) were either epileptics being treated at a seizure clinic or brain surgery patients being seen at a neurosurgery clinic. The remaining 24 *Ss* (the *Non-Brain-Damaged-Adult* subgroup) were either part of the general medical population of a hospital or were nonpatient *Ss*.

The age and a measure of intelligence were obtained for all *Ss*. In addition, the length of institutionalization was obtained for the Defective group. The *Ss* in the Defective group were routinely tested on the Stanford-Binet upon admission to the training school; all other *Ss* were tested for IQ just prior to presenting them with the CPT. The intelligence measure used for the Child group was the Stanford-Binet vocabulary, while the *Ss* in the Adult group were tested on either the Full Scale Wechsler-Bellevue, the Verbal Scale of the Wechsler-Bellevue, or the 20-minute modification of the Otis Self-administering Test of Mental Ability, Form B. Approximately equal numbers of *Ss* in each of the Adult subgroups were tested on each of the three measures of intelligence. The average ages and IQs of the several groups are presented in Table 1. The mean age of the Non-Brain-Damaged-Child subgroup was significantly higher than that of the Brain-Damaged-

Table 1

Mean Ages and IQs of the Subjects

Group	*N*	Mean age	*SD* age	Mean IQ	*SD* IQ
Defectives					
Brain-damaged	29	27.3	8.1	63.3	7.8
Non-brain-damaged	43	25.1	7.8	66.8	7.5
Children					
Brain-damaged	19	9.0**	3.0	101.9	28.2
Non-brain-damaged	26	12.3**	2.9	104.6	21.2
Adults					
Brain-damaged	25	31.7	12.1	90.5*	15.2
Non-brain-damaged	25	31.5	13.9	100.0*	15.0

* Indicates a significant difference ($p < .05$).
** Indicates a significant difference ($p < .01$).

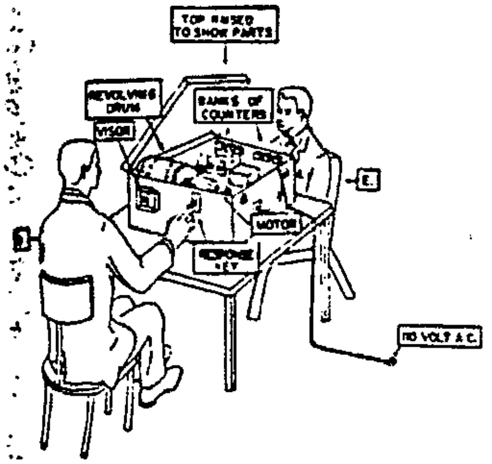


Fig. 1. Semischematic diagram showing the Continuous Performance Test apparatus, a subject and examiner in the positions they would occupy during a test.

subgroup ($p < .01$). Also, the IQ of the Non-Brain-Damaged-Adult subgroup was significantly higher than that of the Brain-Damaged-Adult subgroup ($p < .05$). There were no other significant differences between the various paired subgroups with respect to age or IQ. The mean length of institutionalization for the Brain-Damaged-Defective subgroup was 11.4 years, for the Non-Brain-Damaged-Defective subgroup, 9.3 years. This difference was not significant.

Apparatus. The testing instrument is illustrated in Fig. 1. It consisted essentially of a revolving drum on which two series, each of 31 letters, were mounted side by side.

One series served for the *X* task, the other for *AX* task. The drum revolved slowly, approximately twice a minute, in a boxlike case equipped with a visor through which the *S* was required to look. The *E* shifted the task from *X* to *AX* by changing the horizontal position of the visor. The letters were illuminated briefly one at a time. The illumination was provided by a bank of five 0.5-watt neon bulbs mounted inside the case, fired by means of the discharging of an 8-microfarad condenser. The *S* performed by pressing a response key when certain "correct" letters appeared. In the *X* task, every *X* was a correct letter; in the *AX* task, every *X* following an *A* was a correct letter. The letters came into view at approximately 0.92-sec. intervals. A response was scored correct if the response key was pressed within 0.69 seconds after the brief illumination of the correct letter. Following this interval, there was a period of approximately 0.23 seconds before the next letter was illuminated. Correct and incorrect responses were recorded automatically on banks of counters facing the back of the machine where *E* sat. The *X* task was always presented first and included eight *X*'s in the series of thirty-one letters. The responses occurring over the first two revolutions of the drum served as practice and were not counted in the *S*'s score. The *Ss* were then given either a 5-minute test (10 revolutions) in which there were 80 possible correct *X* responses, or a 10-minute test (20 revolutions) in which there were 160 possible correct responses. Following a 2-minute rest period, the visor was moved by *E* so that the *AX* series was in view. There were six correct *AX* sequences among the thirty-one letters of the *AX* series. As with the *X* task, the responses occurring during the first two drum revolutions served as practice. Then, the *Ss* received either a 5-minute test in which there were 60 possible correct *AX* responses, or a 10-minute test, in which there were 120 possible correct.

Procedure. No attempt was made to keep the instructions completely identical from *S* to *S*; the typical form, however, did not vary greatly from this content:

"Do you see this viewer? When the machine starts, you will see letters appearing one at a time. Your job is to press this key every time you see an *X*. Don't press it for any other letter and always press it when you see an *X*. Press the key down for a bit when you see an *X*; don't hit it too quickly. Try holding your eyes different distances from the viewer until you find the one that is most comfortable for you. Remember, press the key every time you see an *X*, but not for any other letter." Those *Ss* who had difficulty in understanding were given further instruction of the same sort.

With the *X* task and the two-minute rest period completed, the *S* was instructed as follows:

"Your job in this next part is again to press the key when you see an *X*, but this time, only if the *X* follows right after an *A*. That is, when you see the series *A—X—* press the key on the *X*. Don't press the key for any other letter following an *A* except *X*, and always press it then. Remember now, when you see the letter *A*, get set; if an *X* comes right after it, press the key." Additional instruction was given to those who had difficulty in understanding.

Except for the children, all *Ss* were tested for 10 minutes on each of the two tasks. With children, the time was reduced to 5 minutes on each task, and in addition, the instructions were modified to some extent, presenting the task as a game. For the children under the age of 10, pretesting was carried out to make certain that they could discriminate letters, and to see whether they could follow the instructions by requiring answers to two special test sheets. one containing *X*s among other letters, the other, *A*s and *X*s among other letters. Those who failed this pretest were not tested further. No *S* was tested who could not perform the minimal motor tasks of holding his head up unsupported and pressing the response key. Since the testing was conducted at many different clinics and hospitals, it was not feasible to control exactly the amount of illumination in the testing room. The ordinary lighting of the room was used without any modification; care was taken however, to insure that there was no glare on the visor of the testing apparatus.

Reliability. Two measures of the reliability of the CPT were made, using *S*s of the Defective group. Odd-even reliability was determined from the scores of 21 non-brain-damaged *S*s, by comparing the subtotal of correct responses on the 10 even revolutions of the drum with the subtotal on the 10 odd revolutions. The Pearson *r* for the *X* task was 0.88, for the *AX* task 0.86. To determine test-retest reliability, 21 brain-damaged and 22 non-brain-damaged *S*s were retested from four to seven weeks after the initial testing. The test-retest *r* for the number of correct responses of the brain-damaged group on *X* was 0.90, for *AX*, 0.79. The test-retest *r* for the number of correct responses of the non-brain-damaged group on *X* was 0.88, for *AX*, 0.74.

CPT scores. The *X* and *AX* responses of each *S* were scored in two ways. The absolute percentage correct was defined as the number of correct responses the *S* made divided by the number of correct responses (× 100) possible in the time allowed. The relative percentage correct was defined as the number of correct responses (× 100) the *S* made divided by his number of attempts (both correct and incorrect) during the time allowed. In the event that the two scores differed markedly, depending on the *S*'s pattern of responding, the two measures provided a more complete description of his performance than would either of the two taken singly.

Results

CPT scores. Table 2 presents the means and standard deviations of the absolute percentage of possible correct responses on the *X* and *AX* tasks for the six subgroups. Since there was a significant age difference between the Child subgroups and a significant IQ difference between the Adult subgroups (Table 1), the differences between the paired subgroup means on *X* and *AX* in these two groups were evaluated by means of an analysis of covariance. The means of the Child group in Table 2 have been adjusted for regression on age and the means of the Adult group have been adjusted for regression on IQ. The covariance procedure was corrected for heterogeneity of variance between the subgroups when the analysis required it. The mean differences in the Defective group were evaluated by means of the *t* test.

The differences between the subgroup means were highly significant on both the *X* and

Table 2

Adjusted Means and Standard Deviations for Child and Adult Groups, Means and Standard Deviations for Defective Groups, *X* and *AX* Absolute Percentage Correct Measures

	Brain-damaged			Non-brain-damaged			
Group	*N*	Adj. mean	*SD*	*N*	Adj. mean	*SD*	*p* diff.
Children							
X	19	41.61	28.39	26	82.57	14.65	<.0005
AX	19	46.46	28.20	26	79.41	17.44	<.0005
Adults							
X	25	78.53	28.26	25	88.00	17.21	>.30
AX	23	61.40	26.66	23	78.77	15.33	<.01
Defectives							
X	29	46.44	32.02	43	67.86	25.62	<.01
AX	29	32.30	25.79	43	54.74	20.02	<.0005

Table 3

Significance Levels of Differences Between X and AX Scores, for Relative and Absolute Measures of Performance, All Groups

Group	Brain-damaged Absolute	Brain-damaged Relative	Non-brain-damaged Absolute	Non-brain-damaged Relative
Children	>.70	>.30	>.50	>.70
Adults	<.001	<.01	<.01	<.05
Defectives	<.01	<.001	<.001	<.001

Table 4

Significance Levels of Differences Between Adjusted AX Mean Scores of Brain-Damaged and Non-Brain-Damaged Subgroups, All Groups

Group	Performance measure: Absolute	Performance measure: Relative
Children	>.20	<.05
Adults	<.01	>.05
Defectives	<.05	<.05

AX tasks in the Defective group ($p < .01$ and $p < .005$, respectively) and also in the Child group ($p < .0005$ in each case). In the Adult group, the brain-damaged subgroup was significantly inferior to its controls on AX ($p < .01$), but not on X ($p > .30$).

When the *relative* percentage scores of the paired Child and Adult subgroups were compared by means of an analysis of covariance, and those of the Defective group by means of the t test, the paired subgroup differences were all in the same direction and of approximately the same magnitude as those reported in Table 2. The levels of significance were in no case less than the corresponding levels reported in Table 2. The only difference was that the *relative* scores on the X task discriminated significantly ($p < .05$) between the Adult subgroups, whereas the absolute scores on the X task (Table 2) did not.

In order to determine if the AX task was in fact more difficult than the X task, the mean X score was compared with the mean AX score for each of the six subgroups. Table 3 presents the significance levels from the t test of the differences between X and AX scores for both the relative and absolute measures of performance in all six subgroups. Each of the Defective and Adult subgroups performed significantly worse (at $p < .05$ or less) on the AX than on the X task. With the children, none of the differences between X and AX was significant. To determine if the brain-damaged subgroups had relatively more difficulty on AX (i.e., independent of their performance on X) than the non-brain-damaged subgroups, an analysis of covariance was run on the AX scores of the paired brain-damaged and non-brain-damaged subgroups with X scores as the covariate. Table 4, which presents the results of the analysis, shows that in terms of the absolute measures

Table 5

Percentages of Correct Identification of Diagnostic Category by CPT Scores, All Groups

Subgroup	Defectives: Brain-damaged	Defectives: Non-brain-damaged	Children: Brain-damaged	Children: Non-brain-damaged	Adults: Brain-damaged	Adults: Non-brain-damaged
X	29	43	19	26	25	25
$X-R$	75.9***	67.4**	84.2***	76.9***	68.0*	68.0*
$X-A$	62.0	58.1	89.5***	76.9***	64.0	64.0
$AX-R$	72.4**	65.1*	84.2***	76.9***	66.7*	66.7*
$AX-A$	75.9***	67.4**	84.2***	76.9***	65.0†	65.0†

* $p < .05$.
** $p < .01$.
*** $p < .005$.
† Based on 23 subjects.
Note.—p-values refer to the probability that the percentage differs from chance (50%). Nonstarred percentages do not differ significantly from chance.

of performance, the Brain-Damaged-Adult and -Defective subgroups had significantly greater difficulty on the *AX* task than their respective nondamaged subgroups, even when the paired subgroups were adjusted to equal levels of performance on the *X* task. Similarly, in terms of the relative measures of performance, the Brain-Damaged-Child and -Defective subgroups had significantly greater difficulty on *AX* than their respective non-damaged subgroups.

Correct identification of subgroup. Table 5 presents the percentages of correct identification of diagnostic category of all the subgroups on all measures yielded by the CPT. These were determined for each subgroup on each CPT measure from the over-all median of scores of a group on that measure. For example, considering the *X* scores in the Defective group, the percentage of brain-damaged *Ss* scoring *below* the combined Defective group median defines the percentage of correct identification of brain-damaged defectives on the *X* task. Similarly, the percent-

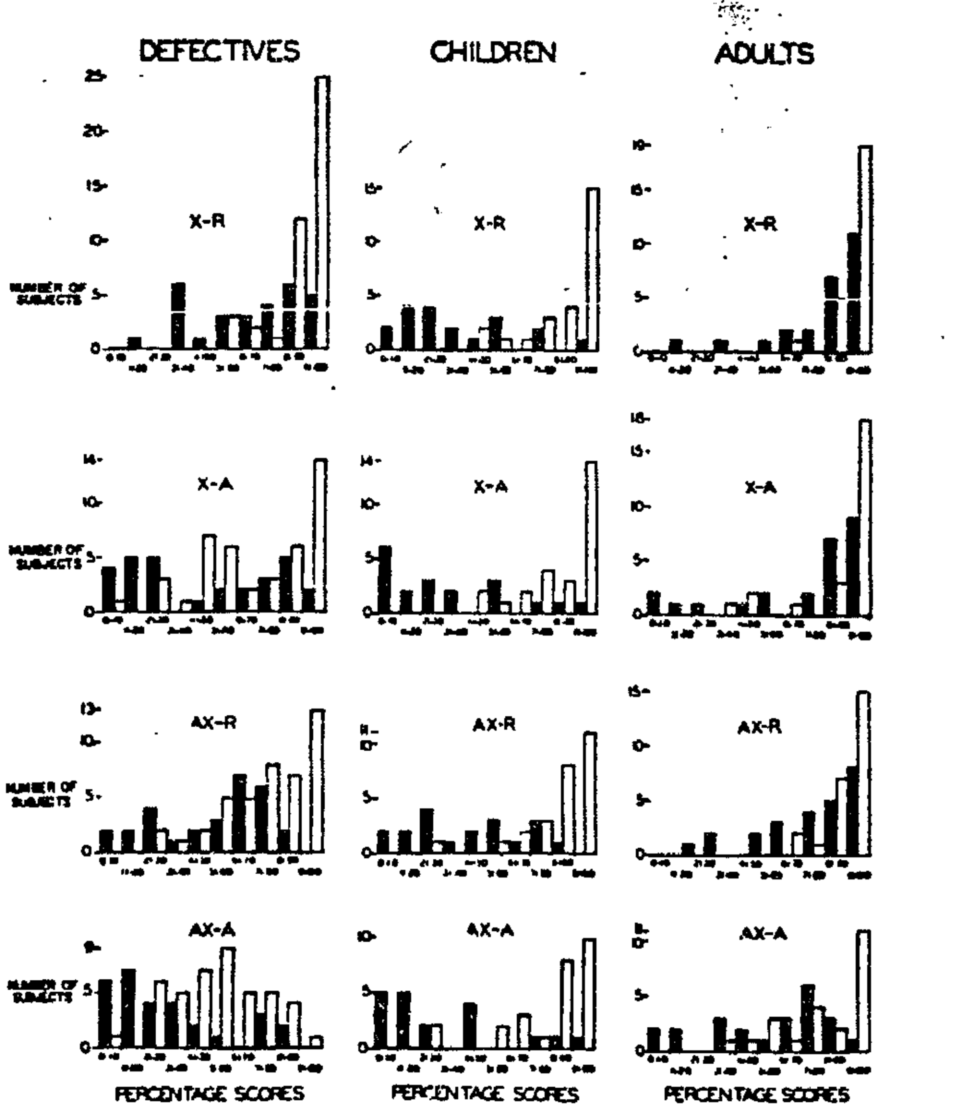


Fig. 2. Frequency distributions of subjects by percentage correct score intervals, all groups, all CPT measures. Black columns represent frequencies of brain-damaged subjects, white columns, frequencies of non-brain-damaged controls.

age of non-brain-damaged defectives scoring *above* the over-all median defines the percentage of correct identification of non-brain-damaged defectives on the *X* task. These percentages were computed to provide an estimate of the potential usefulness of the several scores of the CPT as diagnostic indicators in various clinical groups. In Table 5, *X-R* and *AX-R* refer to the relative measures of performance on these tasks; *X-A* and *AX-A* refer to the absolute measures of performance. The relative measures provided more significant percentages of correct identification than did the absolute measures; all of the six *X-R* and all of the six *AX-R* subgroup percentages were significant at $p < .05$. The Child group had more significant percentages of correct identification than any other (eight of eight) all of which were significant at $p < .005$. There were equal numbers of significant percentages of correct identification among the non-brain-damaged and brain-damaged subgroups; however, these percentages tended to be higher among the brain-damaged subgroups.

Figure 2 presents for each group and CPT measure, the distribution by percentage score of brain-damaged and non-brain-damaged *Ss* (uncorrected for age or IQ differences between subgroups). Figure 2 emphasizes in general the great range in scores achieved by brain-damaged *Ss* as compared with non-brain-damaged *Ss*. The scores of the non-brain-damaged *Ss* are usually concentrated near the upper end of the distribution; in fact, the modal score for non-brain-damaged *Ss* is in the 91 to 100% interval in every case except for the *AX-A* measure in the Defective group. In contrast with this, the scores of the brain-damaged *Ss* span the entire distribution, or the entire distribution excepting one score interval, in every case.

Discussion

The results of this study indicate that brain-damaged individuals perform poorly relative to non-brain-damaged controls on a task requiring continuous attention. The results also indicate that on a task designed to require even more sustained attention the brain-damaged individuals perform even more poorly. The score distributions of the various subgroups (Figure 2) and the percentages of individuals correctly identified by diagnostic category (Table 5) suggest that the CPT is sufficiently sensitive to the effects of brain damage so that it might prove useful clinically. The data of Table 5 suggest that the relative measures are more discriminating than the absolute measures, that the test is somewhat more reliable in reflecting the presence than the absence of brain damage and that it is most discriminating when applied to the children of this study.

These findings are of course preliminary. Use of the CPT in the clinic must await standardization on a variety of diagnostic groups, to ascertain whether brain pathology only and which types of brain pathology lead to deficient performance. For example, some data which have not been reported here indicate that schizophrenics perform no differently from normals of comparable age and IQ. More information concerning other groups is required. Standardization on a variety of diagnostic groups is also necessary in order to have a reliable diagnostic score for each subgroup. Thus, in the present study, cutoff points which included most brain-damaged children missed many brain-damaged adults (see Figure 2).

Diagnostic considerations aside, the question remains as to whether the impaired performance of the brain-damaged *Ss* on the CPT is in fact due to an impairment of attention, whether in the form of momentary lapses, or of a lowering of a general level. Clearly, the available information does not rule out alternate interpretations of the deficit. To make certain that it is attention that is impaired would require more extensive information about the *S*'s test behavior than the present instrument provides. It is clear, furthermore, that the available information does not permit the conclusion that the impaired CPT performance is indeed related to hypersynchrony as evident in the EEG. Nonetheless, the results of this study do lend support to these hypotheses suggesting that it might now be profitable to make a more direct test of the original formulation. This test would be provided by simultaneously measuring CPT performance and brain activity to determine whether errors on the CPT are

all CPT

15

coincident in time with hypersynchronous patterns on the EEG and whether there is a correlation between degree of deficit on the CPT and degree of hypersynchrony.

Summary

1. The Continuous Performance Test, a procedure for the detection and study of brain damage in humans, is described.

2. Three groups of *S*s, each including a brain-damaged and a non-brain-damaged subgroup, were tested on this procedure.

3. The brain-damaged subgroups were significantly inferior to their non-brain-damaged controls on the measures yielded by the CPT, and these differences were increased when the difficulty of the task was increased.

4. The CPT is sufficiently reliable and yields sufficiently large differences between subgroups to suggest that it might ultimately prove useful as a clinical instrument for the diagnosis of brain damage.

5. An interpretation of the inferior performance of the brain-damaged *S*s was offered in terms of impairment in attention or alertness and suggestions were made about future research relating cerebral events and CPT performance.

Received January 13, 1956.

References

1. Hebb, D. O. *The organization of behavior.* New York: Wiley, 1949.
2. Levin, S., Greenblatt, M., Healey, M. H., & Solomon, H. C. Electroencephalographic and clinical effect of prefrontal lobotomy, with consideration of postlobotomy convulsive seizures. In M. Greenblatt, R. Arnot, & H. C. Solomon (Eds.), *Studies in lobotomy.* New York: Grune & Stratton, 1950. Pp. 400–427.
3. Lindsley, D. B. Electroencephalography. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: Ronald. 1944. Pp. 1033–1103.
4. Mettler, F. A. (Ed.) *Psychosurgical problems.* New York: Blakiston, 1952.
5. Mettler, F. A. (Ed.) *Selective partial ablation of the frontal cortex.* New York: Hoeber 1949
6. Partridge, M. *Pre-frontal leucotomy.* Springfield, Ill.: Charles C Thomas, 1950.
7. Petrie, A. *Personality and the frontal lobes.* New York: Blakiston, 1952.
8. Schwab, R. S. *Electroencephalography in clinical practice.* Philadelphia: Saunders, 1951.

16

*Measures of Cognitive Functioning Appropriate for Use in Pediatric Psychopharmacological Research Studies

James M. Swanson, Ph.D.[1]

Introduction

Six types of cognitive tasks, each of which has been used previously in several studies and by several investigators, were selected for description in this paper:

1. Continuous Performance Tasks
2. Choice Reaction Time Tasks
3. Single-Trial Recall Tasks
4. Matching-to-Sample Tasks
5. Paired-Associate Learning Tasks
6. Analogue Classroom Tests

These cognitive tests may be considered to be ordered on a dimension of complexity or difficulty (from "low-level" to "high-level" cognitive tasks, according to Weiss & Laties, 1962), or on a dimension of effortfulness (from "automatic" to "effortful," according to Hasher & Zacks, 1979). Thus, each measures a different aspect of attention or cognition. Each type is available in a computerized and non-computerized version.

The Six Classes of Tasks

1. The Continuous Performance Task (CPT)

The CPT is a monitoring task in which the subject watches an experimenter-controlled sequence of stimuli (e.g., letters) to detect an infrequently occurring target. A "vigilance decrement" in performance occurs within relatively short (5 to 30 min) periods of time. Several versions of the CPT are outlined in Table 2. Detailed descriptions of these CPTs are provided by Nuechterlein (1983). This test has been used to document attentional deficits in hyperactive children (e.g., Sykes et al., 1971, 1973) and children at risk for schizophrenia (Nuechterlein, 1983; Rutschmann et al., 1977). The CPT has a long and distinguished history of use with children, dating back to Rosvold et al. (1956); auditory versions, requiring little equipment (e.g., a tape recorder), to visual versions requiring special purpose projectors and computers are available; it is sensitive to medication effects. However, most versions present a "low-level" task, and the more sophisticated versions may be required to measure psychopharmacological effects on effortful cognitive processes (Nuechterlein, 1983). A related perceptual test, the Span of Apprehension test, is described by Asarnow (1984).

2. Choice Reaction Time Tasks

Choice reaction time tasks have been used to study stages of human information processing (Callaway, 1983; Sanders, 1983; Sternberg, 1969; Welford, 1960). The Sternberg (1966) and Shiffrin and Schneider (1977) scanning tasks, which address four stages of information processing (stimulus encoding, memory comparison, response selection, and response execution), use digits or letters as stimuli. Memory load and display load are the major independent variables, and the dependent measures are reaction time (RT) and error rate. The slope and intercept of a linear RT-load function are interpreted and used to estimate the time required by component stages (e.g., memory comparison time). Sprague and Sleator (1977) used a version of this task with large display loads (up to 15 items), and found that the optimal dose for performance (RT and errors) on it were lower than for subjective rating of behavior. Sergeant and Scholten (1983) and Sergeant (in press) have shown clear differences between normal and ADD/H children in the intercept of the RT-memory load function, but no slope differences, and Callaway

[1]Child Development Center, University of California, Irvine, CA 92717.

*PLEASE NOTE: All references in this paper are presented in a single reference section to be found on pages 1111-1124.

Table 1

Cognitive Tasks Recommended for Use in Pediatric Psychopharmacological Studies

Tests	Source	Selected References
Continuous Performance Tests	Dr. Keith Nuechterlein UCLA-Neuropsychiatric Institute Box 18 760 Westwood Plaza Los Angeles, CA 90024	Nuechterlein, 1983 Sykes et al., 1971, 1973 Rosvold et al., 1956
Choice Reaction Time Tests	Dr. Joe Sergeant Laboratory for Experimental Clinical Psychology Turfsingel 46 Groningen, The Netherlands	Sergeant & Scholten, 1983 Sprague & Sleator, 1977 Sternberg, 1969
Single-Trial Recall Tests	Dr. Herbert Weingartner Lab. Psych. & Psychopath. National Institute of Mental Health Bldg. 31, Room 4C-35 9000 Rockville Pike Bethesda, MD 20205	Weingartner et al., 1980 Rapoport et al., 1978 Craik & Lockhart, 1972
Matching-to-Sample Tests	Dr. Robert Sprague Institute Child Development & Behavior University of Illinois 51 Gerty Drive Champaign IL 61820	Sprague, 1984 Campbell et al., 1971 Kagan et al., 1964
Paired Associate Learning Tests	Dr. James Swanson Child Development Center University of California 19262 Jamboree Road Irvine, CA 92717	Swanson et al., 1983 Gittelman-Klein & Klein, 1975 Conners et al., 1964
Analogue Classroom Tests	Dr. William Pelham Dept. of Psychology Florida State University Tallahassee, FL 32306	Pelham et al., 1985 Stephens et al., 1984 Douglas et al., in press

(1983) and Klorman et al. (in press) have used evoked potential latencies as well as RT to show that methylphenidate does not affect the slope of the RT-memory load function. These findings suggest that ADD/H children have functional, not structural deficits associated with the memory comparison process, and that methylphenidate affects post-evaluative processes, perhaps at the motor selection and execution stages of processing.

3. Single-Trial Recall Tasks

Cognitive psychologists (e.g., Craik & Lockhart, 1972) have developed tests to study specific input (encoding) and output (retrieval) processes of human memory. Lists of words are used as stimuli; the words are typically presented one at a time to the subject, who then attempts to recall them in any order. Independent variables related to encoding and retrieval are established by requiring subjects to process the words in different ways during encoding (e.g., semantically or acoustically), and by guiding the subject with cues (in both encoding and retrieval phases, in one phase but not the other, or in neither phase). Both recall and recognition tests are given, to probe for the availability (recognition) as well as the accessibility (recall) of information in memory (Craik & Tulving, 1975). The studies of Rapoport et al. (1978) and Weingartner et al. (1980) demonstrate the use of this test for assessing the encoding and retrieval strategies of hyperactive children, and the effect of medication on recall. These tests focus on the processing of episodic information in short-term or working memory (one-time recall), and typically do not require "learning" or reorganization of information in semantic memory.

Table 2

Computerized and Non-Computerized Tests Recommended for Use in Pediatric Psychopharmacological Studies

Tests	Computer	Investigator
Computerized		
CPT-Degraded Stimulus	IBM PC	Nuechterlein, Lindgren
CPT-Playing Card	Commodore	Taylor, Lindgren
CPT-AX	Apple IIe	Garfinkel
Span of Apprehension	Apple IIe	Asarnow
Recall	Atari, Apple	Sandman, Swanson
Memory Scanning	Apple II+e,c	Swanson
Pursuit Tracking	Apple II+e,c	Swanson
Matching-to-Sample	Apple IIe	Sprague
PAL Picture-Letter	Apple IIe	Deutsch
PAL Word-Word	Apple II+e,c	Swanson
Arithmetic Tests	Apple II+e,c	Swanson
Non-Computerized	Equipment	Reference
CPT-Auditory	Tape recorder	Campbell et al., 1971
Etch-a-Sketch Tracking	Etch-a-Sketch	Humphries, 1979
Recall Free-Cued	Word lists	Swanson, 1985
Recall Categories	Word lists	Kagan, 1966
Recall Encoding	Word lists	Weingartner et al., 1980
Display Scanning	Letter lists	Neisser, 1967
MFF-20	Notebook	Douglas, in press
PAL Picture-Letter	Slides	Swanson, 1983
PAL Word-Word	Word lists	Douglas, in press
Analogue Spelling	Word lists	Pelham et al., 1985

4. Matching-to-Sample Tasks

The Matching Familiar Figures (MFF) test (Kagan, 1966; Kagan et al., 1964) is a simple version of this test in which a standard or target picture is presented, along with a set of similar pictures which contains one that is identical. The subject's task is to find the exact match. The usual dependent measure is the time-to-first response (a measure of reflection time) and the number of errors made before a match is found. In some studies (e.g., Campbell et al., 1971) stimulant medication has produced a pattern of more reflective responding (longer latencies and fewer errors). In other studies of medication (imipramine and methylphenidate), latency has increased, but errors have not decreased (e.g., Rapoport et al., 1974). Flintoff et al. (1982) used computer-generated forms and recorded eye movements in a Kagan-like MFF test. They found that stimulant medication produced an increase in systematic comparisons of the variants to the standard (an increased selective attention), but no improvement in performance as measured by latency to first responses or errors. Sprague (1984) has developed a sophisticated, computerized version of the matching-to-sample task designed to measure individual differences in strategies as well as levels of performance. It utilizes letter-like stimuli (Gibson figures) which can be systematically transformed to vary difficulty, and computerized monitoring of the subject's selection of stimuli to inspect. Sprague (1984) has shown that the inspection ratio (the number of times each stimulus window is opened per trial) is sensitive to the beneficial effects of low doses, and shows the behavioral toxic effects of high doses in the same way as his earlier memory scanning test (Sprague & Sleator, 1977).

5. Paired-Associate Learning Tasks

In the paired-associate learning (PAL) test, stimulus-response pairs of common, well-known items (e.g., picture-letter pairs, or word-word pairs) are used. The subject's task is to learn the arbitrarily assigned response to each stimulus item, through rote repetition and rehearsal. Each trial consists of a new study-test-feedback sequence, and error scores are calculated for each trial.

Typically, subjects are required to continue the task until reaching a criterion of perfect performance. In one of the earliest series of studies of the effects of stimulant drugs on hyperactive children (Conners, 1973; Conners et al., 1964), and in one of the most recent series of studies (Douglas et al., in press), the PAL test was found to be a good test for distinguishing hyperactive from normal children and for monitoring the effects of stimulant medication. The PAL test has been used by Swanson and Kinsbourne (1976; Swanson, in press-b; Swanson et al., 1978) in more than a dozen studies, and by at least 10 other investigators in studies of ADD/H children (*see* Swanson, in press-b). The paired-associate learning test is a "high level" or "effortful" test, and because of task-specific effects of stimulant medication, lower doses may produce indications of "behavioral toxicity" on this, like on other effortful tests (*see* Sprague, 1984), than on "low level" or "automatic" tests (*see* Rapport et al., in press; Swanson, in press-a). It is one of the best tests for showing differences between ADD/H and normal subjects (Douglas, in press). Medication-induced changes in performance on it are correlated with behavioral changes reported by teachers and psychiatrists (Gittelman-Klein & Klein, 1975). However, some ADD/H children who show clear behavioral improvement may show no improvement or impairment on the PAL test (Swanson et al., 1983).

6. Analogue Classroom Tests

Recently, tests of reading, spelling, and arithmetic have been used to document the effects of medication on academic productivity in the classroom. Pelham et al. (in press; Stephens et al., 1984) used an analogue spelling test, in addition to reading and math tests, to monitor academic production and medication effects in ADD/H subjects. Douglas et al. (in press) have used arithmetic tests to show short-term effects of medication on ADD/H children in the classroom. These tests border on being achievement and psychometric tests, and are for that reason somewhat out of the realm of this paper on cognitive tests for use in pediatric psychopharmacology. However, the analogue spelling test is a laboratory-like classroom test (Pelham et al., in press)

Summary

Table 1 summarizes the six types of cognitive tasks recommended for use in pediatric psychopharmacological studies. These tests are designed to measure different aspects of attention and cognition, but all have been shown to be sensitive to the effects of medications in children, and useful in documenting deficits in clinical groups of children. In general, the "low level" or automatic tests are more sensitive than the "high level" or effortful tests for documenting effects of medication, but the "high level" or effortful tests are more specific than the "low level" or automatic tests for distinguishing groups or subtypes.

The studies by Sprague and Sleator (1977) and Rapoport et al. (1978) are recommended as two of the most important studies in pediatric psychopharmacology of the past decade. The articles by Conners (1973), Gittelman-Klein and Klein (1975), Callaway (1983), and Douglas et al. (in press) are highly recommended and should be useful to anyone planning a study of cognition in pediatric psychopharmacology. Computerized batteries of tests are available from Keith Nuechterlein, (U.C.L.A.) Scott Lindgren (Univ. of Iowa), Barry Garfinkel (Univ. of Minnesota), and James Swanson (U.C., Irvine).

*The Computerized Continuous Performance Test

C. Keith Conners, Ph.D.[1]

The CPT has been used in many studies as one operational definition of "attentional" dysfunction in children. The test consists of a series of visual or auditory stimuli, some of which are designated as targets to which a button-press is required. Usually the test includes an easy version in which a single target is designated, and a more difficult version in which the same stimulus can be a target or nontarget depending upon whether it is preceded by another particular stimulus (such as an "X" when preceded by a "B"). What the test actually measures has never been adequately studied, and investigators are cautioned that such concepts as "sustained attention" and "attention span" are inappropriate for the task when used over brief periods. Generally, the CPT records errors of omission and commission, but not decline in vigilance over time (though there is no obstacle to extending the task and recording such changes).

Unfortunately, no standard version exists. Parameters which need to be considered are total number of stimuli, total test duration, interstimulus interval (ISI), probability of target occurrence, stimulus duration, and method for calculating scores. Omissions and false alarms can be used to calculate signal detection parameters of d' and Beta, the former measuring threshold sensitivity, the latter the response bias or criterion for making a response. Variations in response style, such as cautiousness or impulsivity, affect the speed-accuracy trade-off. While mechanical or computerized methods are desirable, paper-and-pencil methods such as cancellation tasks or "checking" have been standardized (Keogh & Margolis, 1976).

The following are computerized methods which are readily adaptable to a wide age range for drug studies and sample selection.

1. NIMH Version

This early version test has been developed and standardized upon a normative sample. It presents 800 visual stimuli (3/8" high digits) sequentially presented for 100 milliseconds (ms) on a LED display. The task requires the detection of infrequent two-digit sequences (6 followed by 4). The ISI is automatically reduced by 5% following each correct identification and by increasing the ISI by 5% following each error, thus adjusting the rate according to the subject's performance. d' and Beta are calculated (Sostek et al., 1980). The apparatus was designed and built by NIMH, but is not available for sale.

An AX version of the task, currently in use by Rapoport and associates at the NIMH, has a sequence of single letters presented on an LED display. The child pressed a button if X appeared, only if preceded by A. Failure to do so was an omission error; pressing the button for a letter outside the critical sequence was a commission error. Children were instructed to work to maximize the stimulus-presentation rate. A total of 800 digits was presented; there were 32 A-X combinations (4%), 96 A (12%), and 96 X (12%). The mean interstimulus interval for the last 400 stimuli formed a third summary score. The child is usually tested for approximately 10 minutes.

The adaptive version of the task in which the speed is increased or decreased by 5% following a correct or incorrect response, respectively, permits testing over a wide age range using the "sure" task. For younger children (6 and below) the adaptive version often becomes too slow, becoming instead an (extremely) de-

[1]Children's Hospital National Medical Center, Washington, DC 20010.

*PLEASE NOTE: All references in this paper are presented in a single reference section to be found on pages 1111-1124.

21

layed reaction time task. For such subjects, the fixed interval version must be used. While not currently for sale, the NIMH Adaptive AX task is being developed by Michael Gordon and will be available to the public.

2. Klee and Garfinkel (1983)

This version is programmed in PASCAL for the Apple-II computer. Ten letters flash on a video monitor for 130 ms with an ISI of 600 ms. The target is the letter T preceded by the letter S. Ten percent of 500 letters are targets. Errors of omission and commission are calculated.

3. Behavioral Medicine—CPT

This is a program written for the Apple-II computer in BASIC. The stimuli are large alphabetic block letters with a display time of 50 ms. ISI is 1.5 sec. Three hundred trials are presented of which 50 are targets for each task. A training session is optionally usable for standardized practice and instructions. Included with the same diskette is a computerized version of the Sternberg test of short-term memory, and a visual sequential memory test. Available from Behavioral Medicine, Inc., 3904 Cleveland St., Kensington, MD 20895 ($150).

4. Life Sciences Associates Tasks

This is a signal detection task which calculates standard signal detection parameters. The program is written in BASIC and is available for both Apple-II and IBM PC computers. Stimuli are small letters which may be displayed anywhere on the video monitor. Targets are the letter X which occurs as specifiable frequency, duration, and rate amidst distractors (letter Z). The ISI can be varied across a wide range. Available for $49 from Life Sciences Associates, 1 Fenimore Rd., Bayport, NY 11705.

5. The Gordon Diagnostic System (GDS)

This system provides researchers and practitioners with a means of obtaining accurate data on a child's ability to delay and to sustain attention. The portability and ease-of-operation of the GDS has enabled large-scale data collection (Gordon, 1979; McClure & Gordon, 1984).

The GDS allows for the administration of two tasks, the Delay Task and the Vigilance Task, a version of the Continuous Performance Task (Rosvold et al., 1956). A new procedure, the Distractibility Task, is currently being developed to assess the effects of distraction on the child's ability to sustain attention.

A microcomputer software package is available that automatically receives data from the GDS, tabulates the various scores, stores the information, and prints out an interpretive summary of the performance. Included in the report are educational recommendations and suggestions for pharmacotherapy.

Readers interested in further information on the GDS may contact Michael Gordon, Ph.D., at Upstate Medical Center, 750 East Adams Street, Syracuse NY 13210 (Tel. 315-473-8145). Information about purchase of the GDS may be obtained from Clinical Diagnostics, Inc., 300 E. Mineral Ave., Suite 3, Littleton, CO 80122 (Tel. 800-521-4503; 303-795-0438 in Colorado).

All these measure are sold for clinical use.

*Measuring Activity Level in Children

C. Keith Conners, Ph.D.,[1] and Sandra Kronsberg, Ph.D.[1]

"Activity" has become a central concept, both in the understanding of normal development, and of childhood psychopathology. Activity level at a given age, rate of development across the age span, and patterning of activity are parameters which show important relationships with measures of personality, behavioral style, learning, and maladaptive functioning. Activity in man, as in most mammalian species, is a recognizable trait with strong evidence of high heritability (Buss & Plomin, 1975; Buss et al., 1973; Guilford & Zimmerman, 1956; Owen & Sines, 1970; Scarr, 1966; Schoenfeldt, 1968; Thurstone, 1951; Willerman, 1973). It is to be expected that biological variation alone would ensure that some children would be strongly characterized as either under- or overactive, traits which may be predominant in their behavior over much of their life.

However, stable individual differences in activity level are only one part of a more complex picture. Activity level changes over the age span. Almost all methods of measuring activity show a decline of activity level with age, whether by global impression of parents or caretakers, discrete samples of behavior, or mechanical transducers of physical energy. Again, by general principles of behavioral ontogenesis, one can assume that there will be pathologies of development, such as a failure to modulate the level of activity in accordance with normative standards at a particular age. As with pathologies of activity *level*, developmental deviations in activity can simply be the extremes of a normal distribution (in this case, of *rate of change* in activity); or these deviations can be the consequence of a variety of biological insults and environmental stressors.

Activity is not merely a genetic expression of a stable trait, nor simply a parameter of normal development. It is, in fact, not a unidimensional construct, but one which requires specification of environmental variables which act to inhibit, exacerbate, or modulate the behavioral acts of the child—acts which are summed together by one method or another—and labelled "activity." Activity is a resultant vector whose origin may involve the social, educational, and perhaps nutritional environments interacting with the biological substrate of activity. There is great variability in the degree and rate at which raw energy becomes tamed by the socialization process.

It is the environment, most of all, which *patterns* activity level. We may suppose, then, that there will be meaningful categories of environmentally caused pathologies of activity level, which will be best described in terms of an abnormality of *pattern*; of fit between environmental press and temperamentally based activity. Thus, undersocialization is one environmental hazard which may lead to impulse disorders whose expression can resemble hypermotility. Such disorders are, of course, fundamentally different from impulse disorders which derive from the *inability to apply restraint* in the context of normally adequate socialization. Another type of "hyperactivity" is the profound curiosity which sometimes drives the talented child to continually challenge the bounds of boredom, perhaps along the way incurring the displeasure of some adult who finds him "hyperactive." Nor is the underactivity of the dreamy genius the same as the hypoactivity of the forlorn child. These are distinctions which can be made only as part of a pattern of findings, and for this task the human observer far surpasses mere quantitative recordings of changes in physical energy by means of the motion actometer or counts of

[1]Children's Hospital National Medical Center, Washington, DC 20010.

*PLEASE NOTE: All references in this paper are presented in a single reference section to be found on pages 1111-1124.

small units of behavior. Some authors speak of this as *motivated activity*, and relate it to the goals and purposes of the actors within a social context.

It is not surprising that there should be disagreements among different operations for assessing activity level. The human observer may think he or she is as sensitive as the pendulum-driven watch spring which records changes in acceleration. The observer is really sensitive to different dimensions; to the meanings, goals, proprieties, and social codes which constrain behavior. The observer can detect when an event is a reward, a punishment, or a coercion, and grasp the nature of physical or social pain which drives and constrains particular behaviors. Whether the observer is the actor or the other, the patterned appropriateness of activity is best judged by them. When the mother says her child is overactive, she may mean, "For me, in this house, at this time of my life, with my level of patience, and my expectations of behavior, this kid is driving me nuts." Excellent, then, as a measure of social appropriateness and expectation, such judgments may be meaningless with regard to quantitative measures of energy expenditure.

These measures make *independent* contributions to the prediction of behavior. It is perfectly appropriate to ask parents whether their child is up and down at the dinner table a lot, but it is cultural norms which prescribe when a certain level of ambulation during dinner is too much, and which also govern the sanctions that shape the expected behavior. The absolute level of energy consumption is only dimly related to many of these social variables. We may well expect to find, then, that correlations among measures of socially defined *inappropriate* activity and counting methods ("meter reading") of physical energy are low. Barkley and Ullman (1975) found that activity measures tended to correlate with each other, but the relationships were small in magnitude, and differed according to the type of sample. In general, though, hyperactives tended to show a global hyperactivity which generalized across settings, but this was not true for nonhyperactive clinic patients and controls.

We believe, then, that there is a role for quantitative measurement of activity level as a trait whose interactions with drugs is quite meaningful, both as independent and dependent variable. There is also a role for self or other judgments of risk-taking, drive, sensation seeking, and curiosity behaviors which need to be considered equally carefully. Lack of restraint, or inhibitory capability may not be the same thing as innately high drive level towards activity. The rubric of "activity" subsumes many distinguishable notions.

Researchers in psychopharmacology need to be aware of the properties of the various measurement methods and make a judicious choice based upon the requirements of the specific problem and setting. The study of activity is, unfortunately, a methodologic minefield. (Though slightly outdated, one of the best reviews of methodologic issues in activity level research is the comprehensive chapter by Cromwell et al., 1963.)

Rating and Self-Report

Ratings are the method of choice when cost and feasibility are severely limiting factors, and when what is required is an estimate of the fit between the temperamental trait of activity level and the family. Little will be said here regarding these methods since they are discussed elsewhere in this volume.

The Werry-Weiss-Peters checklist (WWP) has been successfully employed as a drug-sensitive measure in a number of studies (Werry, 1968; Werry et al., 1966). This scale records the impressions of parents regarding a child's activity in a number of settings, such as mealtime, television watching, play, sleep, and social settings. The scale has been factor analyzed (Routh et al., 1974), and there is little agreement between activity levels displayed in one home situation and those in another. Rated activity was found to show an almost perfect linear decrease in scores across the age range from 3 to 9 in normal children. Interparent agreement ranged from 0.16 to 0.58 (median = 0.33), and sex was nonsignificant as a variable. Although a similar monotonic decrease in activity was found for direct observation in a playroom, there was no significant relationship between the two types of measures. A similar finding was reported by Barkley and Ullman (1975). These findings imply that one should use this scale not in place of other methods, but as a complement to them. In the absence of more sophisticated (and costly) techniques,

24

it provides a straightforward assessment of activity in the home setting which makes it attractive in psychopharmacologic studies.

The Conners teacher and parent rating scales each have a "hyperactivity" factor which has proven drug-sensitive and diagnostically useful. Along with the 10-item abbreviated scale these methods have been extensively studied, and their ease of use makes them irresistible as "quickie" measures which allow comparison with a large normative database and many previous drug studies. But extensive use has also documented the variable and inconsistent relationship of such ratings to other methods, and caution is recommended in exclusively relying upon ratings since they are sensitive to context, informant, practice effects, age, and sex.

Buss et al. (1980) have provided a Q-sort method for descriptions of children at the preschool level, including energy, restlessness, fidgetiness, caution, aggressiveness, and assertiveness, which shows good validity with actometer measures.

The concept of "impulsiveness" has a close relationship to activity level. Eysenck et al. (1984) have recently published a comprehensive normative study on 1,505 7- to 15-year-old normal children, using a self-report questionnaire. Factor analysis showed that the same factors emerged for boys and girls, and confirmed the expectation that impulsiveness consists of two distinct factors: impulsiveness in action or speech, and venturesomeness (including risk-taking and sensation-seeking). Reliabilities and factor homogeneity are excellent, and the scale may prove useful as a descriptive measure for subject samples, though it seems unlikely to be drug-sensitive given its trait-oriented nature.

Direct Observation

Recording of free-field activity in playroom settings has provided valuable information regarding both normal and hyperactive children (Hutt & Hutt, 1970; Milich, 1984; Pope, 1970; Routh et al., 1974). Generally, these methods require that the room be marked off into grids from which crossings can be calculated. An excellent guide to the development of such methods may be found in Hutt & Hutt (1970).

As described elsewhere in this volume, the observational scale by Abikoff et al. (1980) is recommended as a reliable and valid measure of gross and minor motor behavior within a classroom setting when trained observers are available.

Mechanical Methods

Schulman & Reisman (1959) were among the first to apply the actometer to clinical child studies. The actometer is a wristwatch which has been modified so as to be sensitive to changes in acceleration. The validity of actometer measurement has been seriously questioned by Johnson (1971) who reported that when two actometers were attached to the wrist of someone using a hammer, the more distal actometer had significantly higher readings, suggesting that arm length may be a confounding variable. However, methodological studies have shown that arm length in young children shows no relation to results, and that although recordings from a single actometer are somewhat unreliable, reliability increases rapidly as more than one actometer is measured and several samples are recorded (Eaton, 1983).

Considerable data support the validity of actometer measurement, though there are some conflicting data. Following a suggestion by Bell (1968), Halverson and Waldrop (1973) studied indoor and outdoor play activity in preschoolers, finding significant correlations across settings as well as significant relationships with teachers' ratings of activity. Importantly, the sex differences in activity were detected with the actometer but not with teacher ratings. Sex and age effects were found for staff ratings as well as actometer scores. Actometer and teacher scores were also significantly associated in a study by Buss et al. (1980), who found that actometer measures remained consistent across a 4-year age span. Stevens et al. (1978) also found that actometer scores correlated strongly with mother and trained clinical staff ratings of activity level. Milich (1984) found good intra-subject stability coefficients for actometers over a 2-year period, but only for restricted settings ($r=0.47$) and not free-play.

Kendall and Brophy (1981) examined the interrelationship of teachers' ratings (Conners Hyperactivity factor), a stabilimeter chair, and wrist actometers. The rating factor was significantly but modestly associated with the actometer ($r=0.26$) but not with stabilimeter or behavioral observations. The actometer correlated well with the stabilimeter ($r=0.65$) and the behavior observations ($r=0.53$). The stabilimeter also related well to the behavioral observations ($r=0.58$). It should be noted that the sample was normal school age children ($N=49$), of whom only 10 would have qualified as hyperactive on the Conners scale. Again, age and sex effects upon activity level were noted.

Actometers have been used successfully as a drug response measure by Rapoport et al. (1980), who measured truncal activity in a cognitive testing situation with amphetamine-treated normal and hyperactive boys, and normal adult men. This actometer uses a solid-state memory and allows continuous measurement over time, with readouts at selectable time intervals (Colburn et al., 1976). The sensitivity of this instrument was indicated by the fact that the normal adults showed activity decreases with drug, even though they had low levels to begin with.

One of the most important applications of this actometer was by Porrino et al. (1983a), who used the continously recording actometer over a 1-week period with hyperactives and normal controls. Using recordings from consecutive 1-hour periods, they found that the hyperactives were more active in all situations, including sleep. The breadth of sampling made possible by this method effectively eliminates much of the unreliability which occurs from sampling in short intervals in unrepresentative situations. They found that activity during school was the best overall discriminator of the clinical groups. Importantly, this method allowed the investigators to demonstrate that activity and attentional (CPT) measures were independently contributing to the discrimination between groups, misclassifying different subjects: the motor measure misclassified older hyperactives who tended to move less, while the CPT misclassified younger controls who had difficulty with the CPT. In addition, the measure was highly sensitive to stimulant drug effects, with powerful interaction between situation and effect (Porrino et al., 1983b).

Resources for Actometers

Several sources are available for continuously recording actometers:

1) The NIMH actometers developed by Colburn et al. (1976). These are apparently not directly available but could conceivably be manufactured according to the specifications reported in the original article.

2) Mr. Gary Mathews
KFM Corporation, Inc.
c/o Western Psychiatric Institute and Clinic
3811 O'Hara Street
Pittsburgh, PA 15232
(412 624-2353)

Have the rights to the NIH patent for non-government use.

3) Precision Control Design, Inc.
646 A. Anchors Street
Ft. Walton Beach, FL 32548
(904 224-1923)

Advertises complete monitor and microcomputer readout support.

4) Ambulatory Monitoring, Inc.
731 Saw Mill River Road
Ardsley, NY 10502
(914 693-9232)

5) Vitalog Corporation
1058 California Avenue
Palo Alto, CA 94306
(415) 365-1100)

In addition to the cost of the actometer, a microcomputer is required for displaying and analyzing the data.

Several versions of the wrist accelerometer have been available in the past from Timex. However, Timex no longer markets these. They are presently available from:

Kaulin & Willis
282 Watertown Road
Middlebury, CT 06762

The model 100 is a wristwatch which is sensitive to motion along the 6 to 12 o'clock vector. They sell at a cost of $130, with reductions in price for larger quantities. A more useful version for children, the Model 101 motion recorder, is identical to the Model 100 except that the recorder hands cannot be set. This renders the recorder tamper proof from unsupervised children (who generally will wish to try and fiddle with the hands).

To those researchers embarking upon the mechanical measurement of activity level, some cautionary notes are indicated. First, most of the recorders need to be independently calibrated, and in the case of the watch varieties, this may need to be done often. Second, equip-

26

ment breakdown is frequent and somewhat unpredictable. The electronic versions are particularly sensitive to environmental conditions; and showers, fights, and miscellaneous mishaps of childhood render their unsupervised use both costly and frustrating.

→ See page 310

Review of Educational Research
Winter 1977, Vol. 47, No. 1, Pp. 293-318

The Measurement of the Hyperactive Syndrome in Children

Jonathan Sandoval

University of California, Davis

The hyperactive syndrome in school-aged children has become one of today's major mental health problems. Estimates of the prevalence of the syndrome indicate that from 5% to 20% of the elementary school population are affected (Cantwell, Note 1). Evidence is beginning to accrue that links hyperactivity in childhood with later disorders in adolescence and adulthood (Menkes, Rowe, & Menkes, 1967). In school and at home children with hyperactive behavior are extraordinarily difficult to manage and demand an inordinant amount of a parent's and a teacher's time. Although many children may be treated successfully with stimulant medication, this practice has been hotly debated in the press and in professional publications (Grinspoon & Singer, 1973).

Over the last fifteen years an extensive body of studies on the topic of hyperactivity has accumulated, most of which have either described the condition or have evaluated the effectiveness of various treatment regimens for helping children control and change their behavior. Research will undoubtedly continue in this area as new treatments are devised, e.g., diet regulation (Conners, Goyette, Southwick, Lees, & Andrulonis, 1976), and as new medications come onto the market. Other research trends are the prediction of response to medication and the determination of subgroups of hyperactive children on the basis of etiology or other diagnostic information.

I wish to thank Carl Spring and Dana Sassone for their assistance in reviewing the manuscript and for their generous suggestions.

In much of the research to date, there has been an effort to use objective measures of different aspects of the behavior of hyperactive children, both to explore the nature of the syndrome and to evaluate changes in children which may have come about as a result of treatment interventions. In many instances, investigators have approached the selection of measures from either an exceedingly narrow perspective or have used what may be termed a shotgun approach. Typically, they have either selected one measure such as activity level, ignoring other aspects of the hyperactive condition such as impulsivity or short attention span; or they have administered a large battery of tests with little thought to the dimensions of hyperactivity being assessed. This review is an examination of the instruments that have been used to measure different aspects of the hyperactive syndrome in children. The objective has been to collect and order information on the measurement of the hyperactivity syndrome in children so that a second generation of research may yield more comparable and complete information to guide practice in educational and medical settings.

Some Caveats

Before proceeding, I believe it is useful to point out some of the difficulties encountered in the execution of the review. The greatest problem has been the wide variety of subjects who have been called hyperactive, hyperkinetic, minimally brain dysfunctional, or learning disabled. The diagnosis of a hyperactive condition is most often made on the basis of reports of hyperactive behavior, not on the basis of hard, unequivocal medical or laboratory findings (Sandoval, Lambert, & Yandell, 1976). As a consequence, there is little way to know if the subjects in one study are equivalent to those in another. Most often, unless the focus is specifically on the epileptic or the retarded, subjects are described as having hyperactive behavior yet have no other medical condition which would give rise to the behavior. Within these boundaries the range of severity and types of behavior is immense. In addition, in different studies subjects vary with respect to socioeconomic status, ethnic background, sex, and age, to mention the most obvious dimensions. In this review, I have grouped together studies in which subjects may have been quite different. I have, however, excluded from the review studies on epileptic and retarded children, for the most part, as well as studies on learning disabled children who were not specifically termed hyperactive.

The reader must also recognize that few investigators of hyperactivity have concerned themselves with the measurement instruments they have used. The bulk of the literature concerns

evaluations of medications for use with hyperactive children. Because of this focus, extensive information on reliability and validity is not present in research reports. In many cases the instruments are used because they are a part of a standard assessment battery given to each child admitted to a clinic, and were not especially selected as instruments to study.

Dimensions of the Hyperactivity Syndrome

The hyperactive syndrome consists of the following pattern of symptoms: high activity level, low attention span, impulsivity, aggressive behavior, excitability, and emotional lability (Cantwell, Note 1). In addition many children have learning difficulties, although others do not. These traits are most commonly listed in clinical descriptions of hyperactive children (Lambert, Windmiller, Sandoval, & Moore, 1976) and are the most usual child characteristics assessed in clinical trials of medication with hyperactive children. Activity level is the most straightforward attribute to measure because hyperactivity in children implies excessive and uncontrolled physical movement. The attentiveness-distractibility dimension is seen by many investigators to be the key factor in describing hyperactive children (e.g. Conners, 1972b; Douglas, 1972). Conners, for one, feels that all other symptoms in hyperactive children can be traced back to a deficit in this area. Impulsivity is often differentiated from distractibility as acting without thinking and has been documented by the Montreal group as being a critical aspect of hyperactive children (see Douglas, 1972). Other investigators have identified poor social relationships, excitability, mood swings, and aggressiveness as being key defining traits of hyperactive children (Conners, 1970).

Organization of the Review

Measures of all these attributes have been used in studies of hyperactive children. The measures themselves can be classified into five types: (1) behavioral ratings, (2) observation schedules, (3) direct physical measurements, (4) simple performance tests, and (5) higher-order cognitive tests. In the following pages, I will review the reliability and validity of instruments of each type as they have been established in a variety of research investigations with hyperactive children.

This review does not exhaust the list of all of the measures that have been used to evaluate hyperactive children. Instead, the most commonly used simple procedures have been emphasized in the review, and procedures using elaborate equipment have been deemphasized.

The Validity Problem

The validity of an instrument for assessing an aspect of hyperactivity in children may be established by its ability to distinguish between normal and hyperactive children, by its ability to distinguish between children taking medication and those receiving a placebo, by its correlations with other concurrent measures, and by its internal dimensional structure. In all cases, the instrument in question must have face validity with respect to one or more of the symptoms of the syndrome.

The first validity criterion is obvious. A measure must discriminate between children labeled hyperactive and children who are not. However, to be a thorough test, the group of hyperactive children must be so categorized on the basis of a score on a measure or on the basis of a completely independent procedure that has nothing to do with the measure being validated. In most instances children are considered hyperactive on the basis of the physician's diagnosis. The fact that an instrument discriminates between two particular groups means little until the instrument can be cross-validated with different groups. Another consideration for validity evidence is that the hyperactive children must be off medication at the time of testing. In many of the studies reviewed, it is not clear that this drug-free condition was obtained.

The second and most common method of establishing validity, that of comparing the scores of medicated children with placebo or nonmedicated hyperactive children, is informative but depends on the assumption that a particular medication is effective in changing hyperactive behavior. The danger here is that most medications are evaluated on the basis of change scores on the instruments. The circularity of definitions may be entrapping.

Aside from the physician's diagnosis of hyperactivity, another form of concurrent validity is the extent to which an instrument correlates with other measures of hyperactivity. This is a useful method of establishing the utility of a measurement technique as long as the second measure has established validity. The fact that two measures of behavior correlate does not mean that either measure necessarily has any relation to hyperactivity unless one measure has been independently validated.

The sensitivity to active medication is just one form of construct validity. But construct validity, always difficult to establish, is even more difficult to establish with the hyperactive syndrome. There is little consensus and less unchallenged research about the condition that permits the prediction of relationships. For example, should hyperactive children perform better on an instrument as they grow older? Several investigators (e.g., Mendelson, Johnson, & Stewart, 1971) have questioned the assumption that children "grow out" of the hyperactive syndrome. Most follow-up studies have not supported this generalization; younger children nevertheless, seem to be scored as more hyperactive than older ones. Most investigators have found many more males to be affected with the condition than females. One might expect sex differences in performance on instruments as a result. If medication affects test performance, one might also expect a variation in test performance with drug dosage. If a drug alters behavior at a low dosage, assuming a ceiling has not been reached, it should show more of an effect at a higher level. This is, of course, a simple view of a complex phenomenon—the working of a drug—and the reasoning may not hold. Given the precautions, various kinds of group differences in performance can add to the evidence concerning an instrument's validity.

Factor-analytic methods yield evidence about the validity of a multidimensional instrument or procedure to the extent that the related items or isolated factors are conceptually meaningful or have face validity. In the case of the hyperactivity syndrome, which is generally considered to have multiple causes, the existence of a number of factors in any complex measure may be expected. Factor-analysis as a technique has subjective aspects so it is important to beware of making strong inferences without information about the population tested and the relative independence of the items rated.

Review of Instruments

Ratings

Behavioral ratings by teachers. Behavioral rating scales are relatively simple and straightforward. It is not surprising that several scales to rate hyperactive children have been developed for use by teachers (Bell, Waldrop, & Weller, 1972; Conners, 1969; Davids, 1971; Greenberg, Deem, & McMahon, 1972; Peterson, 1961).

Conners' Behavior Rating Scale (BRS) is by far the most extensively employed teacher rating scale in research use (Conners, 1969). The scale's 39 items may be summed to obtain a global score (Conners & Eisenberg, 1963; Steinberg, Troshinsky, & Steinberg, 1971), may be grouped into three subscales, classroom behavior, group participation, and attitude toward authority (Conners, Eisenberg, & Barcai, 1967; Conners & Rothschild, 1968; Finnerty, Soltys, & Cole, 1971), or groups of items may be used for subscale "factor" scores of I Aggressivity, II Inattentiveness, III Anxiety, IV Hyperactivity, and V Sociability (Kupietz, Bialer, & Winsberg, 1972; Rapoport, Quinn, Bradbard, Riddle, & Brooks, 1974).[1] Sprague and Werry have normed the factor scoring system on 291 Illinois school children (Sprague, Christensen, & Werry, 1974). Conners has also created a short 10-item form of the scale called the Abbreviated Teacher Rating Scale (ATRS) that is summed to a total (Conners, 1972c; Sprague & Sleator, 1973). The ATRS is made up of many of the same items related to factors I, II, and IV on the BRS, plus two items on mood.

In the longer version, the scale has good test-retest reliability (Conners, 1969). But, test-retest is not a particularly good method of estimating reliability for ratings (see Mischel, 1968). Interrater reliability would have been a better indicator of how consistently the scale may be used.

All of the factors except III, Anxiety, have differentiated between normal children and hyperactive or emotionally disturbed children (Kupietz et al., 1972). The total scores of the BRS (Steinberg et al., 1971) and of the ATRS (Conners, 1972c; Sleator, Von Neuman, & Sprague, 1974) are lower for children receiving the Dexedrine, Cylert, and Ritalin than for children receiving a placebo. In addition, Sleator and her colleagues (1974) found a linear relationship between the score and dosage of Ritalin in their subjects.

When investigators have used the subscale system of scoring, the "classroom behavior," and "attitude toward authority," scores of the BRS have been consistently drug sensitive (Conners et al., 1967; Conners & Rothschild, 1968; Finnerty et al., 1971). The "factor scores" also appear to be drug sensitive to Imipramine, Dexedrine, Ritalin, Cylert, and Chlorpromazine. In six studies (Conners, 1969; Conners, Taylor, Meo, Kurtz, & Fournier, 1972; Kupietz et al., 1972; Rapoport, Abramson, Alexander, & Lott, 1971; Rapoport et al., 1974; Waizer, Hoffman, Polizoes, & Englehardt, 1974), factor IV, Hyperactivity, was drug sensitive in all; factor I, Aggressivity, was sensitive in four studies; factor II, Inattentiveness, was sensitive in four; factor V, Sociability, in two; and factor III, Anxiety, in only one (Conners, 1969).

Although the BRS has been used in conjunction with other measures, intercorrelations are seldom reported. Rapoport et al. (1971), however, have reported that the score on factor IV correlates "playroom activity change" in an observational study. One other investigation has been located shedding light on the BRS's concurrent validity. Saxon and Starnes (Note 2) report low, significant correlation with an actometer reading of arm movement, and even lower, nonsignificant correlations with actometer leg movement and a measure of the amount of time the child was in motion in the playroom. In the same study, the BRS was not correlated with the Davids (1971) scale or the Bell scale (Bell et al., 1972) although these latter two scales correlated .74.

1. Most investigators do not use true factor scores but rather group together items which have a high loading on a given factor on a single scale with no attention to weights.

The factor structure of the BRS seems to be consistent with our notions of hyperactivity except that factor IV includes items reflecting both activity level and impulsivity. In many situations it would be preferable to keep ratings of these two attributes separate. In addition, the factor labeled Anxiety (III) is lacking in validity and adds unnecessarily to the length of the scale. Recently the ATRS rather than the BRT has been used in clinical trials, perhaps because it is shorter and excludes many of the Anxiety and Sociability items. The global score from this rating would probably be a sensitive measure for use in drug trials, but it is too undifferentiated for other uses.

The second rating scale that has good reliability and validity data for hyperactive children is Greenberg's Hyperactivity Rating Scale (HRS) (Blunden, Spring, & Greenberg, 1974; Greenberg, et al., 1972; Spring, Blunden, Greenberg, & Yellin, 1977). The earlier version of this measure consisted of 40 items divided into 10 categories of four items each rated on a 4-point Likert scale. The categories are: Restlessness, Impulsiveness, Distractibility, Low Concentration, Low Perseverance, Irritability, Resentfulness, Cheerfulness, Social Participation, and Verbal Expression. These categories are not indicated on the rating form, and the items appear in a random order. The ten categories when factor analyzed yielded three interpretable factors: I, Hyperactivity, composed of items in the first five categories; II, Hostility, composed of Irritability and Resentfulness category items; and III, Sociability, composed of items in the last three categories (Blunden et al., 1974). This factor analysis was performed on a group of normal kindergarten boys, not a hyperactive sample, and perhaps, as a result, like the Conners BRS, impulsivity is not differentiated in the factor structure from activity level or from attentiveness. Aggressiveness (Hostility) and Sociability, however, are defined by items on the scale, but this last factor has poor interrater reliability.

The earlier version of the HRS was revised by Greenberg and Spring (Spring et al., 1977) to eliminate as much as possible items related to the factors of Hostility and Sociability. The new version consists of eleven categories, each containing three items. New items in the categories of Work Fluctuation, Excitability, Poor Coordination, Fatigue, and Rapid Tempo have been added to the scale; the categories Irritability and Resentfulness have been collapsed to Negativism; and Low Concentration, Cheerfulness, Social Participation, and Verbal Expression have been dropped from the scale with three items related to Social Withdrawal taking their place. Items in the categories Restlessness, Impulsiveness, Distractibility, and Low Perseverance, which contributed most to the Hyperactivity factor were, of course, retained. Spring has collected extensive norms for the new version and, after noting ethnic differences, has reported them for 1,140 white children on each category by sex and grade level. The new version, when it was factor analyzed, did yield the sought-after Hyperactivity factor (51% of the variance) and one other factor labeled Extraversion (15% of the variance). The Interteacher reliability coefficients for six teachers in two classrooms for the Hyperactivity factor (computed by adding together the ratings on Restlessness, Distractibility, Work Fluctuation, Impulsivity, and Excitability) were as low as .15 and as high as .71 for different pairs of teachers. Further evidence for the validity of this version of the scale comes from significant sex differences in ratings on the 11 categories as well as ethnic group and grade-level differences on various categories.

Besides validity evidence from the factor structure, Greenberg's HRS has been shown to be drug sensitive (Yellin, Spring, & Greenberg, Note 3) and, in an observation study (Blunden et al., 1972), the category "impulsivity" correlated with observed impulsivity although other categories were not correlated with parallel observational categories. Data are beginning to accrue on this measure, indicating its utility for use in clinical trials; but for other purposes, like the Conners scale, it may be too undifferentiated to use alone in studying hyperactive children. Although the HRS has not been used as widely as the BRS, it should be considered a contender for future research uses.

One other teacher rating scale has been used in clinical trials of stimulant medication and has sometimes been shown to be sensitive to active medication. This scale, the Peterson-Quay Behavior Problem Checklist (Peterson, 1961; Quay & Peterson, 1967), perhaps because it was not developed specifically to study hyperactive children, has two global factors, Conduct Problem and Personality Problem. The 66 items may be totaled or placed in subscales including conduct, personality, and immaturity. Knights and Hinton (1969) did not find the total drug score to be drug sensitive, but others have found the total score to be correlated with the activity recorder (pedometer) reading (Victor, Halverson, Inoff, & Buczkowski, 1973). Other than these pieces of information, there is very little reliability or validity evidence about the scale for hyperactive children.

32

Behavior ratings by parents. The rating scales that have been developed for use by parents are not promising devices for use in assessing attributes of hyperactivity over time. They have not been shown to have the same degree of reliability or validity as the teacher rating scales. Conners (1970) (Conners, Rothschild, Eisenberg, Schwartz, & Robinson, 1969) has constructed for parents a lengthy symptoms rating scale that has two versions, one of 73 items and the other of 93 items. Conners found two factors in the scale, hyperkinetic symptoms and neurotic symptoms, and has shown that the first of these is drug sensitive to Dexedrine (Conners et al., 1969). In his 1970 paper in *Child Development*, Conners reports the scale as being made up of six factors, I, Aggressive Conduct Disorder, II, Anxious-inhibited, III, Antisocial, IV, Enuresis, V, Psychosomatics, and VI, Anxious-Immature, but in his 1972 paper (Conners et al., 1972) he claimed eight factors, adding an Impulsiveness factor, an Obsessive factor, and a Hyperactivity factor and dropping the Enuresis factor. All but the Obsessive factor of a short form of the scale were influenced by Dexedrine and Cylert treatment (Conners et al., 1972). Arnold and Smeltzer (1974) have also factor analyzed a version of this scale. The attractiveness of this instrument is clouded by the inconsistent descriptions of its nature. Moreover, its length, its unknown reliability, and the technical nature of the wording (Sprague & Sleator, 1963) make it unwieldy.

The most widely used parent rating scale is the Werry-Weiss-Peters Home Activity Rating Scale (HARS) (Werry, Weiss, Douglas, & Martin, 1966; Werry, 1968). The Home Activity Rating scale contains 31 items in seven settings: (1) During meals, (2) Watching television, (3) Doing homework, (4) Playing, (5) Sleeping, (6) Behaving away from home (except school), and (7) Behaving at school.

The HARS was originally designed to be completed by mental health professionals during an interview with a child's parents. Using this method, two long-time colleagues were able to demonstrate an interrater reliability of .90 (Werry et al., 1966). In many studies, however, the parents themselves have completed the scale (e.g., Knights & Hinton, 1969; Rapoport et al., 1971). Routh, Schroeder, and O'Tuama (1974), using a shortened 22-item questionnaire version, obtained median interparent correlations of .33 for 140 pairs of parents of normal children. This figure is comparable to the low interteacher correlations obtained by Spring (Spring et al., 1977). The HARS measure has been shown to be drug sensitive (Conners et al., 1969; Knights & Hinton, 1969; Rapoport et al., 1971; Werry et al., 1966). According to Knights and Hinton, the items most significantly changed under Ritalin were related to distractibility and activity level.

Routh and his group (Routh et al., 1974) factor analyzed their short form of the scale and obtained seven factors that accounted for 67% of the variance. They labeled the factors Television-watching Behavior, Bedtime and Sleeping Habits, Mealtime Behavior, Attention-Getting Behavior, Hyper Verbal Behavior, General Restlessness, and Behavior at Play. It is interesting to note that parents in this study did not view hyperactivity as uniform over all settings for their children. This finding does not coincide with many notions about the consistency of hyperactive behavior across settings. In a separate study, Routh (Note 4) was also able to demonstrate that the parents of hyperactive children rated their children higher than parents of normal children, but found no relationship between ratings and observed hyperactivity in the playroom.

Some studies (e.g., Greenberg et al., 1972; Knights & Hinton, 1969; Millichap, Aymat, Sturgis, Larson, & Egan, 1968; Winsberg, Bialer, Kupietz, & Tobias, 1972; Yellin et al., Note 3) have had parents complete a teacher-rating scale. When this has been done, the parents are able to detect change when their children have been receiving active medication. It would be informative to learn if the same factor structure defines behavioral dimensions when the parents completed the rating scale as when teachers did so. Although the literature generally reports parent rating measures as being useful, it may be the case that many negative results go unreported.

Sleator and Sprague (1974) reported, "Parent behavior ratings did not vary significantly with dosage level nor did they distinguish between drug and placebo conditions. This was true for both school days and weekend days. Parental insensitivity to drug effects has been replicated in all of the authors' studies" (p. 29). In another article, Sprague and Sleator (1973), argued that parents "may not be a reliable source of information as to the effects of the psychotropic drug on the behavior of their children" (p. 730). The items from parent-rating forms may not reflect home behaviors that would be affected by medication, or it may be that the effects of medication are not visible in the home. Parents do not see their children in a structured, demanding situation where they must attend; also they interact with their children after school when the effects of the medication may have worn off. Other than the studies reported above, there are no other reports of the reliability or validity of parent ratings. If they are to be used, the HARS seems to be the most useful.

Criticisms of Ratings

Although ratings are ubiquitous in studies of hyperactivity, they are subject to substantial criticism. One set of criticisms springs from the way ratings are incorporated into experimental designs. In most studies teachers or parents are asked to rate a

child "blind," not knowing whether or not he is in a placebo group or an active medication group. The purpose of concealing the status of the child is to obtain an unbiased rating. But usually the teacher is asked to rate only one child in the class and cannot avoid becoming more sensitive to that child's behavior. As a result, most studies show the "Hawthorne" or "Placebo" effect in the fact that, compared to the initial rating, the placebo group is rated as improved (e.g. Knights & Hinton, 1969). This placebo effect is particularly true of parent ratings (Rapoport et al., 1971). The change in the placebo children may represent a real change attributed to the attention of parents and physician, but it may also result from changes in teacher sensitivity to the rated characteristics. The difference between the placebo group and the active drug group may be statistically significant, but the real difference in terms of meaningful behavior differences may be quite small.

Most of the ratings consist of leading questions (Stewart & Haller, 1975) and cover only symptoms that are clearly negative, none that are positive or positively stated. A better method of collecting rating data would be to ask teachers to rate the entire class or a subgroup of the class, including the hyperactive child, on a scale that contained items related to hyperactivity and items unrelated to hyperactivity.

Ratings may also be improved by providing raters with more concrete descriptors than the usual Likert scale options of "Often," "Somewhat," or "Not at all" (Mischel, 1968). It is not surprising that interrater reliabilities are as low as they are, given the vague terms used in most scales.

Observation Schedules

From time to time investigators have used observational techniques to study hyperactive children, usually to assess the effects of stimulant medication. Observers have most often counted the occurrence of various categories of events. Observations have been made in classrooms (Blunden et al., 1974; Doubros & Daniels, 1966; Patterson, James, Whittier, & Wright, 1965; Sprague, Barnes, & Werry, 1970); and in play settings (Ellis, Witt, Reynolds, & Sprague, 1974; Rapoport et al., 1971; Schleifer, Weiss, Cohen, Elman, Cvejic, & Kruger, 1975; Victor et al., 1973; Whitehead & Clark, 1970) with observers either making frequency counts on the spot or later through viewing a film or videotape (Ellis et al., 1974; Lee & Hutt, 1964; Victor et al., 1973). The observation variables consisted of high inference global ratings by observers in some instances (Rapoport et al., 1971; Victor et al., 1973), but usually the variables were counts of the

33

occurrence of specific low inference behaviors recorded every 5 to 30 seconds.

The rating scales generally are multidimensional and assess several different child attributes. Activity level has been observed by such techniques as dividing a playroom into equal areas, then counting the number of times a child moved across the imaginary grid (Rapoport et al., 1971, 1974; Routh et al., 1974). Other measures of activity level were "Velocity of Movement" (Ellis et al., 1974), "Restlessness" (Blunden et al., 1974), "Number of times up and out of seat" (Schleifer et al., 1975; Sprague et al., 1970), "Whole Body Movement" (Whitehead & Clark, 1970), and "Vigor of Play" (Victor et al., 1973). Attentiveness was measured by counts or ratings of such variables as "Length of Visit to Play Equipment" (Ellis et al., 1974), and "Change of Activity" (Rapoport et al., 1971). Impulsivity has been included only on one observation instrument (Blunden et al., 1974). Mood, excitability, and sociability have been observed by such categories as "Cheerfulness," "Social Participation," "Verbal Expression" (Blunden et al., 1974), and "Withdrawal" (Whitehead & Clark, 1970). Aggression has been measured by counts of "Aggressive Acts" (Schleifer et al.) and "Physical Contact" (Sprague et al., 1970). No single observation instrument has been constructed that measures all hyperactive child attributes, however.

The interrater reliabilities reported have been high (r = .80—1.00) for such low inference behaviors as number of grid crossings, change of activity, aggressive acts, or times child moves away from a work table. Reliabilities were moderately high (71—78% agreement) for moderately inferential judgments such as beginning of "impulsive acts." Highly inferential observations such as "mood" have shown lower interrater agreements (50% agreement or lower).

Two studies attempted to identify hyperactive children from normal children through observations. Schleifer and his colleagues (1975) were unable to observe differences between normal and hyperactive children in a free play setting, but when the preschool children were moved to a structured setting where they were required to play at a table under supervision, the observational measures of activity level and aggressive acts did characterize hyperactive children. Victor and his colleagues (1973) observed elementary school-aged boys in a free play setting but not in a structured situation. Their observation ratings of "hyperactivity" and "vigor of play" did not identify the hyperactive children from controls, although in the same study the activity recorder reading did. Routh et al. (1974), however, were able to establish that grid crossings may be used to differentiate hyperactive from normal children both under instructions to play freely and under instructions to remain in one area.

When observational methods have been used in evaluations of psychoactive drugs, in some studies, observers have detected changes in behavior induced by amphetamine-like drugs, but not tranquilizers (Rapoport et al., 1971; Routh, 1975; Schleifer et al., 1975; Sprague et al., 1970; Whitehead & Clark, 1970). Other studies with Ritalin have not found the observational methods capable of distinguishing between actively medicated and placebo groups (Ellis et al., 1974; Rapoport et al., 1974). The paradox may be resolved in that those observational studies of stimulant-medicated children in structured situations have noted differences, whereas those studies of children in informal situations have not, leading Ellis et al. (1974) and Sroufe (1975) to conclude that stimulants reduce activity level and increase attention capacity in structured situations that demand concentration but do not affect free play behavior.

In some instances, the observational variables have shown good concurrent validity. Activity level observations have correlated with the activity recorder (Victor et al., 1973), attention observations have correlated with teacher ratings of activity level (Rapoport et al., 1971), and impulsivity observations have correlated with teacher ratings of impulsivity, sociability, attention, and activity level (Blunden et al., 1974). In the Blunden study, however, observations of Restlessness, Irritability, Cheerfulness, Social Participation, and Verbal Expression did not correlate with teacher ratings, and in Routh's studies grid crossings did not correlate with parent ratings (Routh et al., 1974; Routh, Note 4).

Observational methods are valuable tools for studying hyperactive children. They are, however, quite costly. Observer training must be extensive to obtain high interrater reliability. Future investigators must be careful to note the context of the observation and should include operational observation categories covering all the attributes of hyperactivity.

Measures of physical activity. Millichap and Fowler (1967) in their methodological review of studies on hyperactivity suggest that direct measures of activity be included in evaluation studies of hyperactive behavior. Direct measures of activity in children have been taken by fitting a room with photoelectric cells (Ellis & Pryer, 1959) or ultrasonic sensors (McFarland, Peacock, & Watson, 1966; Saxon & Starnes, Note 2) by placing a child's desk on a carefully suspended platform (Foshee, 1958) or by placing a radio transmitter in a helmet worn by the child (Davis, Sprague, & Werry, 1969; Herron & Ramsden, 1967). Most of these devices have been used with retarded children, however. Investigators of hyperactive children have commonly used three devices: the actometer, the activity recorder, and the stabilimetric cushion. These are all mechanical devices for sensing movement. They are, respectively, a modified self-winding wristwatch (actometer), a two-dimensional pedometer-type device attached to the child's shirt back (the activity recorder), and a cushion embedded with sensitive microswitches to detect any squirming while seated. The actometer is very reliable when attached to a machine (Saxon & Starnes, Note 2; Schulman, & Reisman, 1959); when attached to a child, however, Johnson (1971) concluded the device was unreliable. Rapoport et al. (1971) noted that the horizontal and vertical measures of the activity recorder correlate .49, which indicates only moderate reliability for this instrument. These measures, as a result, must be regarded as having questionable reliability.

The actometer has been used in three clinical tests of medications by Millichap (Millichap & Boldrey, 1967; Millichap et al., 1968; Millichap & Johnson, 1974) with mixed results. In his earlier study with Boldrey, he found an *increase* in activity in a small number of subjects taking Ritalin compared to subjects taking a placebo. In his second study (Millichap et al., 1968) the instrument did not distinguish between children on medication and the same subjects taking a placebo although scores did seem to distinguish between children on and off medication. In his 1974 study the children most active before treatment with Ritalin were less active on medication, whereas some children (28%) displayed increased activity under Ritalin therapy. In general these subjects who were more active had been less active before treatment. In the 1968 study, however, the most active children at the start did not have different actometer readings under Ritalin.

The activity recorder has been validated with hyperactive children in that it distinguishes between teacher-judged hyperactive and nonhyperactive boys (Victor et al., 1973), but it does not appear to be drug sensitive (Rapoport et al., 1971). The Stabilimetric Cushion not only differentiates between groups (Sykes et al., 1971), but it also seems to be drug sensitive to the stimulant Ritalin although not to the tranquilizer Thioridazine (Sprague et al., 1970; Sprague et al., 1974).

Although these three measures are superficially similar, their use in research settings has been quite different. Both Actometer and Activity Recorder measurements have been made in play situations or over a 24-hour period. The Stabilimetric Cushion, however, has been employed with children engaged in engrossing psychological tests. As Sroufe (1975) has pointed out, it may well be that stimulant medication has the effect of increasing activity in situations demanding little concentration but decreasing activity in settings requiring attention. The results with the three measures might have been similar if they had been used in the same experimental contexts.

Because of the possibility of instrument unreliability, the expense of the equipment, the obtrusive nature of some instruments, and the problem of determining or controlling the context of measurement in a field study, direct measures of physical activity are less attractive than other diagnostic instruments.

Simple Performance Tests

Common visual-motor or performance measures. Four standard perceptual-motor measures have been used in the assessment of hyperactive children: the Bender Visual Motor Gestalt Test, the Frostig Developmental Test of Visual Perception, the Human Figure Drawing Test, and the Porteus Maze Test. Since these measures are often used by psychologists in the schools, they are of particular interest although other performance measures also have been used.

The Bender Gestalt, commonly scored with the Koppitz system, seems to be a measure of both perceptual accuracy and motor coordination. In addition, receiving a low (good) score on the test requires some planning and monitoring of behavior and attention to detail. Of the various aspects of hyperactivity, the Bender score is perhaps loosely related to attention and impulsivity. Most investigators, however, have included it in their test batteries as a measure of motor-coordination. (Others, such as Page, Bernstein, Janicki, & Michelli, 1974, have used the Lincoln-Oseretsky as a measure of motor development.) Large scale studies using the Bender Gestalt have not shown the measure to be sensitive to medication (Conners & Rothschild, 1968; Conners et al., 1969; Conners et al., 1972; Conrad, Dworkin, Shai, & Tobiessen, 1971; Greenberg et al., 1972; Knights & Hinton, 1969; Millichap et al., 1968; Rapoport et al., 1971; Winsberg et al., 1972) or to differentiate normal children from hyperactive children (Palkes & Stewart, 1972). Dykman and his colleagues (Dykman, Ackerman, Peters, & McGrew, 1972), however, were able to demonstrate that the test could differentiate between learning disabled children and normal children but warned that the difference occurs most in younger children who draw the figures "too rapidly, leading to careless reproduction (p. 47)." Evidently the test has value to the extent that it measures impulsivity in some children; however, there are perhaps better measures of impulsivity.

Another test of visual-motor perception and coordination is the Marianne Frostig Developmental Test of Visual Perception. The Frostig yields an overall perceptual quotient based on scores on five subtests: eye-motor coordination, figure-ground discrimination, form constancy, position in space, and spatial relations. Factorial studies of the test indicate that the items are tapping a single ability, global visual perception skills (Mann, 1972). In three out of four instances where the Frostig was used, the perceptual quotient was drug sensitive (Conners, 1972c; Conners et al., 1969; Conrad et al., 1971). In the fourth study (Millichap et al., 1968), the figure-ground subtests, differentiated between active medication and placebo. This subtest was not drug sensitive in the studies by Conrad and Conners and his associates, but tests, Position in Space (Conrad et al., 1971), and Spatial Relations (both Conrad et al., 1971 and Conners et al., 1969) were drug sensitive. (It should be noted that Conrad selected his subjects on the basis of a low Frostig perceptual quotient.) The Frostig seems to be reflecting some aspects of the attributes of hyperactivity. Inspection of the test indicates that attention to detail in the face of distracting stimuli is at the core of the figure ground test, but doing well on all of the tests depends on refraining from responding impulsively.

The Human Figure Drawing test has long been used as a nonverbal measure of intellectual functioning as well as a clinical tool for assessing various intra- and interpersonal dynamics. The test, when scored by the Harris-Goodenough method, gives a child credit for recollecting and including in his drawing various details of a human figure. When the performance on the drawing is given an IQ score, the test was sometimes drug sensitive (Conners, 1971; Millichap et al., 1968) and sometimes not (Greenberg et al., 1972; Page et al., 1974; Rapoport et al., 1971; Rapoport et al., 1974). When the measure is used in clinical trials with hyperactive children, presumably it is used as a measure of attention to detail rather than as an intelligence test. There is no reason to suspect that short-term drug therapy would affect ability per se (Conners, 1972a), although it may affect attention, concentration, and impulsivity, which would lead to improved performance on the measure. This notion is supported by the findings of Palkes and Stewart (1972) who found that hyperactive children scored lower on the Figure Drawing Test than normal subjects when WISC IQ was partialled out. In addition, in one study (Millichap et al., 1968) the drug effect was strongest for children who had initially scored low on this measure, but were otherwise average in intelligence.

Two investigators have departed from the Harris-Goodenough scoring method with mixed results. Conners (1971) found a "communication organ" score to be drug insensitive. Crowe (1972) using a "formal accuracy" score, found the Figure Drawing test could distinguish hyperactive children from normal children and was drug sensitive. Presumably Crowe's results also reflect attention to detail.

36

Although the Human Figure Drawing may be useful in drug studies as a measure of attention rather than intelligence, there are more direct measures of attention available for use in following the progress of hyperactive children. The clinical use of the test will probably continue in the search for new understandings about these difficult children.

The other common standardized performance test used to assess hyperactive children's behavior is the Porteus Maze Test. This performance test consists of a series of mazes graded in difficulty that yield an IQ score (variously known as the quantitative score, the Test Quotient, the Test Age score and the T-score), and a Qualitative or Q-score. Both scores have been used to evaluate hyperactive children. Porteus maintains that the IQ score is a measure of foresight, planning, and the ability to profit from experience. A reviewer (Horn, 1972) describes this score as the "capacity or inclination to sustain attention in the face of the difficulty involved in resolving moderately complex spatial relations" (p. 756). The T-score has been shown to differentiate hyperactive from normal subjects (Spring et al., 1976) and in clinical trials with cerebral stimulants, the T-score is drug sensitive (Conners, Eisenberg, & Sharpe, 1963; Conners & Rothschild, 1968; Conners et al., 1969; Conners et al., 1972; Epstein, Lasagna, Conners, & Rodriguez, 1968).

The Q-score, on the other hand, is generally judged to be "indicative of impulsivity or impetuousness—an inability or unwillingness to refrain from behaving in a hasty, slapdash manner" (Horn, 1972, p. 756). A child may receive a high Q-score because of several different types of performance such as going too rapidly, not following instructions, or general carelessness in motor performance. Some of these behaviors fall into the category impulsiveness, others do not. Thus a high Q-score may reflect various qualitative behaviors in contrast to a high T-score, which is purported to be a measure of planning or attention.

When the Q-score has been used as a variable in clinical drug studies, it has not been found to be drug sensitive (Epstein et al., 1968; Rapoport et al., 1974). Palkes, Stewart, and Kahana (1968) in an interesting study illustrated that hyperactive boys given training in verbal mediation to inhibit impulsive behavior showed improvement as reflected in the Q- and T-scores of the test. This was, however, the only evidence of construct validity discoverable.

In sum, these four common psychological tests seem only indirectly to measure attention and/or impulsivity. They are not, of course, measures of aggression, sociability, activity level, excitability, or mood.

Laboratory measures. More specialized performance measures have been explored in laboratory settings. A group of investigators at Montreal, for example, has looked at measures of automatic functioning that are subject to distraction; the Stroop Color Distraction Test and the Santostefano and Paley Colour Distraction Test. These tests require the subject to read a color-name in the presence of a distracting factor, the color of the ink. Measures from these tests of susceptibility to interference and distractibility did not differentiate hyperactive children from normal controls (Campbell, Douglas, & Morgenstern, 1971; Cohen, Weiss, & Minde, 1972) and were not drug sensitive (Campbell et al., 1971; Waizer et al., 1974). In the Campbell study, however, the hyperactive subjects did make more errors of commission and made fewer self-corrections on the Color Distraction test than normal children did and improved in these respects on drugs. Perhaps this later index of what may be impulsivity from the tests has more utility than the indices of attention-distractibility. It may be that hyperactive children have short attention spans but are no more subject to interference *while they are attending* than normal children (Douglas, 1972).

Another set of laboratory performance measures used by a number of investigators center around speed, reaction time, and errors on simple tasks which demand the child's attention for various lengths of time (Campbell et al., 1971; Cohen, Douglas, & Morgenstern, 1971; Conners & Rothschild, 1968; Jacobs, 1972; Spring, Greenberg, Scott, & Hopwood, 1973; Sykes, Douglas, & Morgenstern, 1972, 1973; Sykes, Douglas, Weiss, & Minde, 1971). Anderson (Note 5) for example, has adapted a vigilance task similar to one used in human factors research, where a subject views a pair of colored lights flashing every two seconds and responds only when a certain combination appears. These tasks seem to be reliable (Test-retest r = .50—.87, Sykes et al., 1973), but fatigue and practice-effect must be taken into account when they are used. Most of these measures differentiate between normal and hyperactive children, particularly on tasks which demand complex decisions over an extended period of time. In these situations, hyperactive subjects do not differ in speed of response but make somewhat fewer correct responses and make significantly more errors of a type which results from impulsive responding (Sykes et al., 1971, 1972, 1973). Medication seems to improve functioning on these tasks when actively medicated children are compared with placebo children. Sroufe (1975) points out that the drug effects can be explained by a quicker deterioration in the placebo group's performance over time. As a result these instruments may be considered a measure of sustained attention.

Most of these techniques use sophisticated or elaborate equipment not suited to field studies. They are, however, important

...easures of attention. Spring, Yellin, and Greenberg (1976) have ...uggested two alternative measures on which hyperactive chil... dren differ from normal children, and which could be used in field studies. They are extended versions of the Digit Span Forward and Coding subtests of the Wechsler Intelligence Test for Children. As they are longer than the versions on the standard WISC, each becomes a measure of short-term memory and sustained attention, but one, Coding, is drug sensitive, and the other, Digit Span, is not. Spring argues that Digit Span may tap an aspect of attention and memory that is an enduring characteristic of hyperactive children whether or not they are treated with medication.

Higher-Order Cognitive Test Performance

Numerous studies have related hyperactive behavior to cognitive test performance. Although other tests have been used, I will concentrate on the Wechsler Intelligence Scale for Children (WISC), the Matching Familiar Figures Test (MFF) and the Embedded Figures Test (EFT). The WISC, of course, is a widely used measure of intellectual functioning that is routinely administered to children experiencing difficulty in school or entering child guidance clinics. As a result, it has been included in many studies on medication, not as a measure of intelligence, usually, but as an index of improvement in attention and concentration. The MFF and the EFT are measures of the cognitive styles of Reflection-Impulsivity and Field Dependence-Independence.

Because of its common use as a clinical instrument, it is surprising that more investigators have not sought subtest patterns on the WISC to use in differentiating hyperactive children from other populations as well as examining the Performance IQ score, the Verbal IQ test score, and the Full Scale IQ score, for different groups. Keogh and her associates (Keogh, Wetter, McGinty, & Donlon, 1973) divided the WISC subtests into a Verbal-Comprehension score made up of the Information, Vocabulary, and Comprehension subtests; an Analytic-Field-Approach score, made up of Object Assembly, Block Design, and Picture Completion; and an Attentional-Concentration score, made up of Arithmetic, Digit Span, and Coding subtest. Hyperactive subject's scores were depressed on the Attentional-Concentration score, but not on the other scores (cf. Spring et al., 1976).

In another study examining subtest patterns, Palkes and Stewart (1972) found that Similarities, Picture Completion, and Mazes scores of a control group were significantly higher than those of a group of hyperactive children. The IQ scores were al- higher.

In clinical trials research, the WISC has been used often but with mixed results. The Full Scale and Verbal IQ measures seem to fluctuate as functions of Performance IQ. Children under medication scored better than children under placebo on the performance scale score in some investigations (Epstein et al., 1968; Knights & Hinton, 1969; Page et al., 1974), but not in others (Conners et al., 1969, Conners et al., 1972; Conrad et al., 1971; Finnerty et al., 1971). When the performance score did improve, Knights and Hinton established that only Picture Completion, Block Design, and Coding had changed, whereas all subtests were higher (no significance tests reported) in the Epstein study. The Coding test here emerged as being drug sensitive in at least one study and was related to an Attentional-Concentration factor in Keogh's division. Here is additional support for using a Coding task as a measure of attention. The use of the WISC itself, however, as a measure of attention and concentration is neither appropriate nor justified, considering the availability of other measures.

The MFF Test was designed by its author Jerome Kagan (Kagan, Rosman, Day, Albert, & Phillips, 1964) to measure the cognitive style impulsivity-reflectivity. An extensive literature exists establishing the construct validity of this instrument. In four studies (Campbell et al., 1971; Cohen et al., 1972; Rapoport et al., 1974; Schleifer, et al., 1975) the MFF has been shown to differentiate hyperactive children of various ages from control children. In addition, in all four studies the impulsivity score has proved to be influenced by active psychostimulants. This is one measure of impulsivity that is consistently useful for studying hyperactive children.

Field independence-dependence, assessed by the EFT (Witkin, 1959), has also been used in research on hyperactivity. The EFT is said to measure one aspect of impulse control, distractibility. Field independent children, the argument goes (Campbell et al., 1971) are better able to act on a problem (locate a figure) in a confused and distracting context (other lines forming a "ground") than are field-dependent children who are presumably distracted. Hence the EFT is used as a measure of attention-distractibility. There is some evidence that hyperactive children are more field dependent than control subjects, but a child's scores on the test does not seem to be influenced by active medication (Campbell et al., 1971; Cohen et al., 1972; Schleifer et al., 1975; Winsberg et al., 1972). As with other instances where attention is assessed in the presence of distractors, there is little drug effect. The EFT is probably not as unequivocal a measure of

attention-distractibility as other tests, but it remains an option for the investigator.

Conclusion

It is evident that most measures used by investigators to study hyperactive behavior are still in a developmental phase. For the majority of the instruments reviewed, reliability coefficients have not been calculated for subjects who are hyperactive. In addition, validity for the measures has been established primarily by only two methods: ability to discriminate between known groups (hyperactive children and normal children) and sensitivity of the instrument to changes in behavior following medication. Considering the nature of the conclusions being made on the basis of data from these measures, they deserve more study.

Campbell and Fiske (1959), in their important article on convergent and divergent validity, point out considerations that should guide research. It is important that different measures be used in evaluation studies and that the correlations and relationships between instruments be examined with the utmost care. Investigators might well consider using teacher and parent ratings, observations, physical measures, simple performance tests, and higher-order cognitive tests in a single study. At the same time they should select measures of activity level, attention-distractibility, impulsivity, aggressiveness, emotional liability, excitability, and sociability from within these measurement techniques. The quality of future research may be improved by the careful planning of a measurement system which employs different kinds of measures of various attributes.

The instruments reviewed also hold promise for clinical practice, both as aids to diagnosis and as tools for evaluation. Selection of a particular instrument by a clinician might be governed by the range of treatment regimens he or she plans to employ. Physicians using medication would obviously attend to the drug sensitivity of an instrument in deciding its appropriateness, whereas therapists using behavior modification techniques would be more interested in instruments that yield accurate baseline information rather than whether the instrument was drug sensitive. The practitioner must be aware of the problems of reliability and the fact that satisfactory norms have been produced for only a handful of the instruments. Limitations aside, the clinician should seek information from sources other than personal observation and an informal interview of parents.

Reference Notes

1. Cantwell, D. P. *The hyperkinetic syndrome.* Chapter to appear in *Recent advances in child psychiatry*, edited by M. Rutter and L. Hersov.

2. Saxon, S. A., & Starnes, K. D. *Hyperactive children: Rating scale validity, parent training, and punishment effects.* Paper presented at the meeting of the American Psychological Association, Chicago, 1975.
3. Yellin, A. M., Spring, C., & Greenberg, L. M. *Effects of imipramine and methylphenidate on behavior of hyperactive children.* Submitted for publication, 1975.
4. Routh, D. K. *Validity of open field locomotion as a measure of hyperactivity.* Paper presented at the meeting of the American Psychological Association, Chicago, 1975.
5. Anderson, R. P. *The vigilance task: A computer based technique for assessing hyperkinesis.* Paper presented at the meeting of the American Psychological Association, Chicago, 1975.

References

Arnold, E., & Smeltzer, D. J. Behavior checklist factor analysis for children and adolescents. *Archives of General Psychiatry*, 1974, *30*, 799-804.

Bell, R., Waldrop, M., & Weller, G. A rating system for the assessment of hyperactive and withdrawn children in preschool samples. *American Journal of Orthopsychiatry*, 1972, *42*, 23-24.

Blunden, D., Spring, C., & Greenberg, L. M. Validation of the classroom behavior inventory. *Journal of Consulting and Clinical Psychology*, 1974, *42*. 84-88.

Campbell, D., & Fiske, D. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 1959, *56*, 81-105.

Campbell, S. B., Douglas, V. I., & Morgenstern, G. Cognitive styles in hyperactive children and the effect of methylphenidate. *Journal of Child Psychology and Psychiatry*, 1971, *12*, 55-67.

Cohen, N., Douglas, V., & Morgenstern, G. The effect of methylphenidate on attentive behavior and autonomic activity in hyperactive children. *Psychopharmacologia*, 1971, *22*, 282-294.

Cohen, N., Weiss, G., & Minde, K. Cognitive styles in adolescents previously diagnosed as hyperactive. *Journal of Child Psychology and Psychiatry*, 1972, *13*, 203-210.

Conners, C. K. A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry*, 1969, *126*, 884-888.

Conners, C. K. Symptom patterns in hyperkinetic, neurotic, and normal children. *Child Development*, 1970, 667-682.

Conners, C. K. The effect of stimulant drugs on human figure drawings in children with minimal brain dysfunction. *Psychopharmacologia*, 1971, *19*, 329-333.

Conners, C. K. Pharmacotherapy of psychopathology in children. In H. C. Quay & J. L. Werry (Eds.), *Psychopathological disorders of childhood.* New York: Wiley, 1972. (a)

Conners, C. K. Stimulant drugs and cortical evoked responses in learning and behavior disorders in children. In W. L. Smith (Ed.), *Drugs, development and cerebral function.* Springfield: Thomas, 1972. (b)

Conners, C. K. Symposium: Behavior modification by drugs: II Psychological effects of stimulant drugs in children with minimal brain dysfunction. *Pediatrics*, 1972, *49*, 702-706. (c)

Conners, C. K., & Eisenberg, L. The effects of methylphenidate on symptomatology and learning in disturbed children. *American Journal of Psychiatry*, 1963, *120*, 458-469.

Conners, C. K., Eisenberg, L., & Barcai, A. Effect of dextroamphetamine on children. *Archives of General Psychiatry*, 1967, *17*, 478-485.

Conners, C. K., Eisenberg, L., & Sharpe, L. Effects of methylphenidate (Ritalin) on paired associate learning and Porteus Maze performance in emotionally disturbed children. *Journal of Consulting Psychology*, 1964, *28*, 14-22.

Conners, C. K., Goyette, C. H., Southwick, D. A., Lees, J. J., & Andrulonis, P. A. Food additives and hyperkinesis: A controlled, double-blind experiment. *Pediatrics*, 1976, *58*, 154-166.

Conners, C. K., & Rothschild, G. H. Drugs and learning in children. In J. Hellmuth (Ed.), *Learning disorders* (Vol. 3). Seattle: Special Child Publications, 1968.

Conners, C. K., Rothschild, G., Eisenberg, L., Schwartz, L. S., & Robinson, E. Dextroamphetamine sulfate in children with learning disorders. *Archives of General Psychiatry*, 1969, *21*, 182-190.

Conners, C. K., Taylor, E., Meo, G., Kurtz, M. A., & Fournier, M. Magnesium pemoline and dextroamphetamine: A controlled study in children with minimal brain dysfunction. *Psychopharmacologia*, 1972, *26*, 321-336.

Conrad, W. G., Dworkin, E. S., Shai, A., & Tobiessen, J. E. Effects of amphetamine therapy and prescriptive tutoring on the behavior and achievement of lower class hyperactive children. *Journal of Learning Disabilities*, 1971, *4*, 45-53.

Crowe, P. Aspects of body image in children with the symptoms of hyperkinesis (Doctoral dissertation, George Washington University, 1972). *Dissertation Abstracts International*, 1972, *33*(4), 1785B.

Davids, A. An objective instrument for assessing hyperkinesis in children. *Journal of Learning Disabilities*, 1971, *4*, 499-501.

Davis, K. V., Sprague, R. L., & Werry, J. S. Stereotyped behavior and activity level in severe retardates: The effect of drugs. *American Journal of Mental Deficiency*, 1969, *73*, 721-727.

Doubros, S., & Daniels, G. An experimental approach to the reduction of overactive behavior. *Behavior Research and Therapy*, 1966, *4*, 251-258.

Douglas, V. I. Stop, look and listen: The problem of sustained attention and impulse control in hyperactive and normal children. *Canadian Journal of Behavioural Science*, 1972, *4*, 259-282.

Dykman, R. A., Ackerman, P. T., Peters, J. E., & McGrew, J. Psychological tests. In C. K. Conners (Ed.), *Clinical use of stimulant drugs in children.* Amsterdam: Excerpta Medica, 1974.

Ellis, M. J., Witt, P. A., Reynolds, R., & Sprague, R. L. Methylphenidate and the activity of hyperactives in the informal setting. *Child Development*, 1974, *45*, 217-220.

Ellis, N. R., & Pryer, R. Quantification of gross bodily activity in children with severe neuropathology. *American Journal of Mental Deficiency*, 1959, *63*, 1034-1037.

Epstein, L. C., Lasagna, L., Conners, C. K., & Rodriguez, A. Correlation of dextroamphetamine excretion and drug response in hyperkinetic children. *Journal of Nervous and Mental Disease*, 1968, *145*, 136-146.

Finnerty, R. J., Soltys, J. J., & Cole, J. O. The use of D-amphetamine with hyperactive children. *Psychopharmacologia*, 1971, *21*, 302-308.

Foshee, J. Studies in activity level: I. Simple and complex task performance in defectives. *American Journal of Mental Deficiency*, 1958, *62*, 882-886.

Greenberg, L. M., Deem, M. A., & McMahon, S. Effects of dextroamphetamine, chlorpromazine and hydroxyzine on behavior and performance in hyperactive children. *American Journal of Psychiatry*, 1972, *129*, 532-539.

Grinspoon, L., & Singer, S. Amphetamines in the treatment of hyperkinetic children. *Harvard Educational Review*, 1973, *43*, 515-555.

Herron, R., & Ramsden, R. Continuous monitoring of overt human body movement by radio telementry: A brief review. *Perceptual and Motor Skills*, 1967, *24*, 1303-1308.

Horn, J. L. Review of Porteus Maze Test. In O. K. Buros (Ed.), *The seventh mental measurements yearbook.* Highland Park, N.J.: The Gryphon Press, 1972.

Jacobs, N. T. A comparison of hyperactive and normal boys in terms of reaction time, motor time, and decision making time under conditions of increasing task complexity (Doctoral dissertation, University of California, Los Angeles, 1972). *Dissertation Abstracts International*, 1972, *33*(3-A), 1045.

Johnson, C. Hyperactivity and the machine: The actometer. *Child Development*, 1971, *42*, 2105-2110.

Kagan, J., Rosman, B., Day, D., Albert, J., & Phillips, W. Information processing in the child: Significance of analytic and reflective attitudes. *Psychological Monographs*, 1964, *78* (Whole No. 578).

Keogh, B. K., Wetter, J., McGinty, A., & Donlon, G. Functional analysis of WISC performance of learning disordered, hyperactive, and mentally retarded boys. *Psychology in the Schools*, 1973, *10*, 178-180.

Knights, R. M., & Hinton, G. G. The effects of methylphenidate (Ritalin) on the motor skills and behavior of children with learning problems. *Journal of Nervous and Mental Disease*, 1969, *148*, 643-653.

Kupietz, S., Bialer, I., & Winsberg, H. G. A behavior rating scale for assessing improvement in behaviorally deviant children: A preliminary investigation. *American Journal of Psychiatry*, 1972, *128*, 1432-1436.

Lambert, N. M., Windmiller, M., Sandoval, J., & Moore, B. Hyperactive children and the efficacy of psychoactive drugs as a treatment intervention. *Journal of Orthopsychiatry*, 1976, *46*, 335-352.

Lee, D., & Hutt, C. A play-room designed for filming children: A note. *Journal of Child Psychology and Psychiatry*, 1964, *5*, 263-265.

Mann, L. Review of Marianne Frostig developmental test of visual perception. In O. K. Buros (Ed.), *The seventh mental measurements yearbook*. Highland Park, N.J.: The Gryphon Press, 1972.

McFarland, J. N., Peacock, L. J., & Watson, J. A. Mental retardation and activity level in rats and children. *American Journal of Mental Deficiency*, 1966, *71*, 381-386.

Mendelson, W., Johnson, N., & Stewart, M. Hyperactive children as teenagers: A follow-up study. *Journal of Nervous and Mental Disease*, 1971, *153*, 273-279.

Menkes, M., Rowe, J., & Menkes, J. A 25 year follow-up study on the hyperkinetic child with minimal brain dysfunction. *Pediatrics*, 1967, *39*, 393-399.

Millichap, J. G., & Boldrey, E. E. Studies in hyperkinetic behavior II. Laboratory and clinical evaluations of drug treatments. *Neurology*, 1967, *17*, 467-471.

Millichap, J. G., & Fowler, G. W. Treatment of "minimal brain dysfunction" syndrome. *Pediatric Clinics of North America*, 1967, *14*, 767-776.

Millichap, J. G., Aymat, F., Sturgis, L. H., Larsen, K. W., & Egan, R. A. Hyperkinetic behavior. *American Journal of Diseases of Children*, 1968, *116*, 235-244.

Millichap, J. G., & Johnson, F. H. Methylphenidate in hyperkinetic behavior: Relation of response to degree of activity and brain damage. In C. K. Conners (Ed.), *Clinical use of stimulant drugs in children*. Amsterdam: Excerpta Medica, 1974.

Mischel, W. *Personality and assessment*. New York: Wiley, 1968.

Page, J. G., Bernstein, J. E., Janicki, R. S., & Michelli, F. A. A multi-clinic trial of pemoline in childhood hyperkinesis. In C. K. Conners (Ed.), *Clinical use of stimulant drugs in children*. Amsterdam: Excerpta Medica, 1974.

Palkes, H., & Stewart, M. Intellectual ability and performance of hyperactive children. *American Journal of Orthopsychiatry*, 1972, *42*, 35-39.

Palkes, H., Stewart, M., & Kahana, B. Maze performance of hyperactive boys after training in self-directed verbal commands. *Child Development*, 1968, *39*, 817-826.

Patterson, G., James, R., Whittier, J., & Wright, M. A behavior modification technique for the hyperactive child. *Behaviour Research and Therapy*, 1965, *2*, 217-226.

Peterson, D. R. Behavior problems of middle childhood. *Journal of Consulting Psychology*, 1961, *25*, 205-209.

Quay, H. C., & Peterson, D. R. *Manual for the behavior problem checklist*. Champaign, Ill.: University of Illinois, Children's Research Center, 1967.

Rapoport, J., Abramson, A., Alexander, D., & Lott, I. Playroom observation of hyperactive children on medication. *Journal of Child Psychiatry*, 1971, *10*, 524-534.

Rapoport, J., Quinn, P., Bradbard, G., Riddle, K., & Brooks, E. Imipramine and methylphenidate treatments in hyperactive boys. *Archives of General Psychiatry*, 1974, *30*, 789-793.

Routh, D. K. The clinical significance of open field activity in children. *Pediatric Psychology*, 1975, *3*, 3-8.

Routh, D. K., Schroeder, C. S., & O'Tauma, L. A. Development of activity level in children. *Developmental Psychology*, 1974, *10*, 163-168.

Sandoval, J., Lambert, N. M., & Yandell, G. W. Current medical practice with hyperactive children. *American Journal of Orthopsychiatry*, 1976, *46*, 323-334.

Schleifer, M., Weiss, G., Cohen, N., Elman, M., Cvjic, H., & Kruger, E. Hyperactivity in preschoolers and the effect of methylphenidate. *American Journal of Orthopsychiatry*, 1975, *45*, 38-50.

Schulman, J., & Reisman, J. An objective measure of hyperactivity. *American Journal of Mental Deficiency*, 1959, *64*, 455-546.

Sleator, E. K., & Sprague, R. L. Dose effects of stimulants in hyperkinetic children. *Psychopharmacology Bulletin*, 1974, *10*, 29-33.

Sleator, E., Von Neumann, A., & Sprague, R. Hyperactive children, a continuous long-term placebo controlled follow-up. *Journal of the American Medical Association*, 1974, *229*, 316-317.

Sprague, R. L., Barnes, K., & Werry, J. Methylphenidate and thioridazine: Learning reaction time, activity and classroom behavior in disturbed children. *American Journal of Orthopsychiatry*, 1970, *40*, 615-628.

Sprague, R. L., Christensen, D. E., & Werry, J. S. Experimental psychology and stimulant drugs. In C. K. Conners (Ed.), *Clinical use of stimulant drugs in children*. Amsterdam: Excerpta Medica, 1974.

Sprague, R. L., & Sleator, E. K. Effects of psychopharmacologic agents on learning disorders. *Pediatric Clinics of North America*, 1973, *20*, 719-735.

Spring, C., Blunden, D., Greenberg, L. M., & Yellin, A. M. Validity and norms of a hyperactivity rating scale. *Journal of Special Education*, in press, 1977.

Spring, C., Greenberg, L., Scott, J., & Hopwood, J. Reaction time and effect of Ritalin on children with learning problems. *Perceptual and Motor Skills*, 1973, *36*, 75-82.

Spring, C., Yellin, A. M., & Greenberg, L. Effects of imipramine and methylphenidate on perceptual-motor performance of hyperactive children. *Perceptual and Motor Skills*, 1976, *43*, 459-470.

Sroufe, L. A. Drug treatment of children with behavior problems. In F. Horowitz (Ed.), *Review of child development research* (Vol. 4). Chicago: University of Chicago Press, 1975.

Steinberg, G. G., Troshinsky, C., & Steinberg, H. R. Dextroamphetamine—responsive behavior disorder in school children. *American Journal of Psychiatry*, 1971, *128*, 174-179.

Stewart, M. A., & Haller, J. P. Letters to the editor: Hyperactive children. *Journal of the American Medical Association*, 1975, *231*, 134-135.

Sykes, D., Douglas, V., & Morgenstern, G. The effect of methylphenidate (Ritalin) on sustained attention in hyperactive children. *Psychopharmacologia* (Berlin), 1972, *25*, 262-274.

Sykes, D., Douglas, V., & Morgenstern, G. Sustained attention in hyperactive children. *Journal of Child Psychology and Psychiatry*, 1973, *14*, 213-220.

Sykes, D., Douglas, V., Weiss, G., & Minde, K. Attention in hyperactive children and the effect of methylphenidate (Ritalin). *Journal of Child Psychology and Psychiatry*, 1971, *12*, 129-139.

Victor, J. B., Halverson, C. F., Inoff, G., & Buczkowski, H. J. Objective behavior measures of first- and second-grade boys' free play and teachers' ratings on a behavior problem checklist. *Psychology in the Schools*, 1973, *10*, 439-443.

Waizer, J., Hoffman, S. P., Polizoes, P., & Engelhardt, D. M. Outpatient treatment of hyperactive school children with imipramine. *American Journal of Psychiatry*, 1974, *131*, 587-591.

Werry, J. S. Developmental hyperactivity. *Pediatric Clinics of North America*, 1968, *15*, 581-599.

Werry, J., & Quay, H. Observing the classroom behavior of elementary school children. *Exceptional Children*, 1969, *35*, 461-470.

Werry, J., Weiss, G., Douglas, V., & Martin, J. Studies on the hyperactive child III. The effect of chlorpromazine upon behaviour and learning ability. *Journal of American Academy of Child Psychiatry*, 1966, *5*, 292-312.

Whitehead, P. L., & Clark, L. D. Effect of lithium carbonate, placebo, and thioidazine on hyperactive children. *American Journal of Psychiatry*, 1970, *127*, 824-825.

Winsberg, B. G., Bialer, I., Kupietz, S., & Tobias, J. Effects of imipramine and dextroamphetamine on behavior of neuropsychiatrically impaired children. *American Journal of Psychiatry*, 1972, *128*, 1425-1431.

Witkin, H. A. The perception of the upright. *Scientific American*, 1959, *200*, 50-56.

41

Stop, look and listen: The problem of sustained attention and impulse control in hyperactive and normal children*,†

VIRGINIA I. DOUGLAS
McGill University and the Montreal Children's Hospital

ABSTRACT

A research programme was undertaken to investigate the specific disabilities of hyperactive children and to evaluate the effectiveness of one of the stimulant drugs on the measures found to differentiate between hyperactives and normals. It is argued that a core group of symptoms involving inability to sustain attention and to control impulsivity can account for most of the deficits found in the hyperactive group. It also appears that the stimulants exert their main effect on these deficits. Correlational and factor analytic studies suggest that the same constellation of abilities underlies the behaviour of normal children in several areas of cognitive and social functioning.

Over the past several years, together with psychiatric colleagues at the Montreal Children's Hospital,[1] my students and I have been studying children whose presenting symptoms have led to a diagnosis of "hyperactivity" or "hyperkinesis." Typically, the parents and teachers of these children complain that they are "always on the move," seem unable to concentrate, and are overly impulsive. Typically too, the hyperactivity has been present from infancy or very early childhood. I would like to review for you some of the things we have learned and some of the questions we have not yet been able to answer, because I believe that these children present a challenge not just for the clinical psychologist but for the developmental, educational, and physiological psychologist as well.

One of the first difficulties involved in trying to talk about the hyperactive syndrome is the problem of establishing a diagnosis. The literature has relied heavily on clinical descriptions and, to complicate matters further, these descriptions often overlap considerably with those given for

* Presidential Address, presented to the Canadian Psychological Association, St John's, Newfoundland, June 1971.

† This research was supported by Medical Research Council Grant No. MA-1577, Canada Council Research Grants Nos. 60-0664 and 68-1590, and a grant from the Ciba Pharmaceutical Company, Dorval, Quebec.

1 In the early years of the project we worked in close collaboration with Dr G. Weiss and Dr J. Werry of the Department of Psychiatry at the Montreal Children's Hospital. In more recent years Dr K. Minde and Dr G. Morgenstern, also from the Department of Psychiatry, joined our research team. I have found this a most rewarding venture in "interdisciplinary research" and I wish to express my sincere thanks to all of these colleagues.

CANAD. J. BEHAV. SCI./REV. CANAD. SCI. COMP., 4, (4), 1972

children suffering from "minimal brain dysfunction" and "specific learning disabilities," two other diagnostic labels that have recently become popular.

We had hoped that it might be possible to find reasonably objective criteria for making the diagnosis, particularly since many techniques for measuring activity level are now available. However, as we have learned more about hyperactive children, we have come to believe that hyperactivity is only one of a constellation of critical symptoms. I hope to convince you that the inability of these children to sustain attention and to keep impulsive responding under control may be even more important symptoms. Other investigators (Cromwell, Baumeister, & Hawkins, 1963; Werry & Sprague, 1970) have also expressed doubt about relying too heavily on simple measures of activity level. It appears that it may be just as important to consider the quality of the hyperactive child's behaviour as its quantity. Because of his short attention span, he tends to flit from one goal to another. As a result, his behaviour is often fragmented and disorganized and these qualities may contribute to the impression of excessive activity. Thus, we decided that, at least in our early work, we would rely on the reports of parents and teachers and the judgment of experienced clinicians in choosing the subjects to include in our studies.

We have tried to keep our samples as homogenous as possible. In each case, hyperactivity had to be the major presenting complaint and the hyperactivity had to have been present from very early childhood. Both the children's parents and teachers had to report that it was excessive enough to be interfering seriously with adjustment. Children who were retarded were excluded, as were those diagnosed as neurotic or psychotic. We also ruled out children who had histories or clear symptoms of brain damage. Because brain damaged children are often hyperactive (Rosenfeld & Bradley, 1948; Burks, 1960; Levy, 1959; Strauss & Lehtinen, 1947), it has frequently been assumed that hyperactive children must have some kind of subtle brain injury. As we shall see, however, there has been considerable controversy over this assumption. Thus, we felt that it was important to check it out with a group of children who were free of the more obvious signs of brain injury.

Our first goal was to define as accurately as possible the specific disabilities that characterize the hyperactive child. Although a large number of cognitive and learning deficits have been attributed to them, including difficulties with reading, spelling, arithmetic, visual motor problems, and deficiencies in conceptual skills (Burks, 1960; Clements & Peters, 1962; Laufer, Denhoff, & Solomons, 1957; Rosenfeld & Bradley, 1948), there have been disappointingly few studies in which their performance has been compared with that of a matched group of normal controls. As a

result, a kind of clinical folklore has sprung up about the disabilities that are peculiar to hyperactive children and there has been no way to separate fact from fiction.

Another problem has arisen from the kinds of intellectual, cognitive, and personality measures that have been used. Most of the clinical instruments employed by investigators in this area tap highly complex functions which depend upon several poorly defined abilities; many of these measures also have a good deal of error variance. This is true, for example, of the subtests of the Wechsler Intelligence Scale for Children, one of the tests which has been most popular with clinicians (Cohen, 1959). Unfortunately, there is no easy solution to this problem. Our own approach has been highly eclectic. We have chosen tests commonly used by clinical psychologists and we have also borrowed extensively from the batteries that have recently been developed by educational psychologists and special educators in order to diagnose specific learning disabilities. Some of our measures are drawn from the cognitive and personality areas within developmental psychology and others are relatively well-known techniques from the experimental laboratory. We have also gone to the children's schools to find out more about the kinds of problems they are experiencing there. What I would like to do is take you along on an abbreviated version of the search that we have carried out thus far. I will begin with the more global measures and then move on to describe some of our attempts to study these children under more controlled laboratory conditions. Each study that I shall report on has involved 20–50 hyperactive children and an equal number of controls matched with the hyperactives on age, sex, overall I.Q. and socio-economic status. In each study the subjects ranged in age from 6 to 14 years and I.Q.s ranged from approximately 80 to 125 with a mean of approximately 100. It is also important to note that 80 to 90 per cent of the hyperactive children referred to us have been males.

INFORMATION OBTAINED FROM SCHOOLS

One valuable source of information on our children has been their school records and the reports of their teachers. Two honours students, Doba Lewin and Henry Lavigueur, analysed data from teachers' ratings and report cards and from achievement and intelligence tests administered in the school. They found a high incidence of failed grades among the hyperactives. By the time they were 12 years of age, 70 per cent of them had failed one year, and 20 per cent had failed two years. The hyperactive children also had significantly lower grades than the controls on almost all academic subjects. The teachers' ratings which showed the highest differences were concerned mainly with frustration tolerance, concentration, and ability to organize one's own activities. On group administered achievement tests, the hyperactives did significantly more poorly than the normal controls on the Stanford Arithmetic Test and the reading speed score of the Gates Reading Test. They did not do more poorly than the controls, however, on either the comprehension or vocabulary scores of the Reading Test. They also received lower I.Q. scores than the controls on group administered intelligence tests; their scores on these group tests were significantly lower than those obtained on individual intelligence tests administered at the hospital.

Another group of students has taken on the task of collecting observational data on the children's behaviour in the classroom. Sandra Witelson did most of the original work involved in developing the observation schedule and the study of hyperactive-normal comparisons was done by Christine Bradley. She studied two groups of hyperactive children – the mean age of the younger group was 7 years, 9 months and that of the older group was 12 years, 8 months. She obtained the most consistent differences between hyperactives and normals on a coding category that included "purposive (or goal-directed) behaviour not related to the classroom activity." It is important to note that we are not talking here about random or aimless behaviour. The problem seems to be, rather, that the child's goals and the teacher's goals differ. We also observed various kinds of "fidgeting" behaviour but our findings on this category have not differentiated reliably between hyperactives and controls. The differences between the younger and the older hyperactives were also interesting. The younger ones moved around the classroom and vocalized more than their controls. They also showed more disorderly behaviour toward the teacher and attracted more attention from her. The behaviour of the older group on the other hand, was less disrupting: although they also engaged in more purposeful behaviour that was not the classroom activity, in their case this behaviour took place while they were sitting in their seats. They were frequently observed, for example, to be working on the wrong assignment or playing with a toy.

CLINICAL AND LEARNING DISABILITY TESTS

I would like to review next the results we obtained from the clinical and learning disability battery. From time to time the clinical literature has reported that hyperactive children have difficulty on particular subtests of the Wechsler Intelligence Scale for Children (Burks, 1960; Clements & Peters, 1962). However, we found no consistent subtest pattern nor did we find the differences between the verbal and performance sections of the test which have sometimes been described. We have repeatedly found, however, that the hyperactive children show more variability from subtest to subtest. As we shall see later, this kind of variability also typifies their performance on other tasks. The hyperactives have also consistently scored lower than the normals on the Goodenough Harris Draw-a-Person Test, the Bender Visual-Motor Gestalt Test, and the Lincoln-Oseretsky Schedule of Motor Development. They had some difficulty, too, with the Frostig Developmental Test of Visual Perception; this showed up in a significantly lower overall score and a poorer score on the eye-motor co-ordination subtest. This last group of findings seems to suggest that the hyperactive children have unusual difficulties on visual motor tasks and tasks requiring fine and gross motor co-ordination. Similar reports appear in the literature (Clements & Peters, 1962; Laufer, et al., 1957). I would like to point out, however, that besides testing visual-motor ability, all of these tests require the child to perform a task with care and concentration; the Bender is also thought to require an analytic approach. I will return to this point later.

It is important to note that on the remainder of the rather extensive battery used, there were almost no differences between the hyperactive and control subjects. This included an individually administered reading test, a test tapping several different abilities in the language area, and tests of such abilities as auditory discrimination, right-left discrimination and short term memory. In all, there were 41 measures on which we found no differences between the hyperactive and normal children.

AUTOMATED CONCEPT LEARNING TASK

In her doctoral dissertation, Vaira Freibergs (Freibergs, 1965; Freibergs & Douglas, 1969) focused on concept learning in these children. The clinical literature had suggested that they would have difficulty functioning at a conceptual level (Burks, 1960; Rosenfeld & Bradley, 1948). Freibergs used a concept learning apparatus and several concept problems developed by Sonia Osler at Johns Hopkins University (Osler & Fivel, 1961). The machine delivers two pictures at a time. One is an exemplar of a concept such as "flower" or "bird" or a number concept such as "two"; the other is a non-exemplar of the concept. In the case of the bird or flower concept, for example, the non-exemplar may be any one of a variety of objects such as a house, an animal or an aeroplane; in the case of the "two" concept it is another number. The child is told that if he looks at the pictures very carefully he will find that there is something in them, like an idea, which will tell him which one to choose to get a marble as often as possible. He makes his choice by pressing a lever. It is thus possible to observe, step by step, the child's attempts to solve a new problem. A record is kept of the number of errors he makes and the number of trials to criterion.

Performance on these tasks was studied under three conditions: a continuous reinforcement condition in which every correct response was rewarded, a 50 per cent partial reinforcement condition in which every other correct response was rewarded and a delay condition in which the interval between presentation of stimuli was increased from 4 to 8 seconds. Under continuous reinforcement, the hyperactive and normal children reached criterion equally quickly, although on the first problem there were differences in the shapes of their learning curves which suggest possible differences between the groups during the presolution phase. The performance of the hyperactives was highly erratic, showing wide fluctuations from their own mean performance. By the second problem, however, this kind of fluctuation seemed to disappear. In both Freiberg's study and a later study by Sheila Macklin, they also showed excellent transfer from problem to problem. We were struck by the way that learning in the hyperactive children seemed to thrive under the circumstances provided by this teaching device in the continuous reinforcement condition. The stimuli were colourful and interesting and they remained on the screen until the subject made his choice. Perhaps even more important, the machine provided immediate feedback and reinforcement for correct responses.

Freibergs also investigated the children's ability to reverse concepts. After the subjects reached criterion the machine was switched so as to deliver a marble when the *non*-exemplar of the concept appeared. There were no differences between the hyperactive and control children in their ability to reverse. Thus, on this task, we find no evidence to support reports in the clinical literature that hyperactive children tend to perseverate. It should be noted, however, that on this particular task the non-exemplars of the concept provide more variety and novelty than the exemplars; we have other data which suggest that the hyperactives may have a greater pull towards novelty. Doubling the intertrial interval from 4 to 8 seconds did not produce significant differences between the hyperactive and normal subjects although there was a tendency for the hyperactives to do less well.

The most striking difference between the two groups appeared under the 50 per cent partial reinforcement schedule. Although both hyperactives and normals had more difficulty under this condition, the performance of the hyperactive children was much more severely impaired. Sixty-five per cent of them failed to reach a solution within 300 learning trials. There are several possibilities which might explain why the hyperactives did so badly when the ratio of reinforcement was reduced. Unfortunately, it is

45

difficult to sort out the effects of motivational, attentional, and information feedback variables. However, Penny Parry is currently making an attempt at accomplishing this as part of her doctoral thesis. We are hoping that what she learns will provide some valuable insights that may help in training these children.

STUDIES OF ATTENTION

Donald Sykes's doctoral dissertation (Sykes, 1969; Sykes, Douglas, Weiss, and Minde, 1971; Sykes, Douglas, and Morgenstern, 1972) was designed to focus on some critical aspects of attention and followed up earlier work done by Edgar Zurif for his master's degree. There are numerous reports in the literature which suggest that attention is impaired in hyperactive children. Little has been done, however, to define empirically the kinds of attentional problems they demonstrate. Sykes's studies were designed to point up some of the factors that might affect their performance in learning situations.

In one task of choice-reaction-time the child was required to respond to specific stimuli but was warned each time before the stimuli appeared: on this task the hyperactive children had no difficulty. A second task, a serial reaction task, also required responses to particular stimuli but in this case the child had to continue responding over a prolonged period of time (15 minutes). The task was self-paced, that is, each stimulus appeared only after the child had made his response to the previous stimulus. Here the performance of the hyperactive children was somewhat impaired: they made more incorrect responses (responses to wrong stimuli) than the controls but they apparently worked quickly enough to make a similar number of correct responses.

The attention task on which the hyperactive children had most difficulty was a continuous performance task. This was a vigilance task developed by Mirsky and Rosvold (1963). It was experimenter-paced and the subject was required to perform over a 15-minute interval. He had to respond to particular stimuli, for example, the letter X when it was preceded by the letter A. The stimuli arrived automatically on a screen or through earphones. On both the visual and auditory forms of this task the hyperactive children identified fewer of the correct stimuli and also responded more frequently to incorrect ones; several of these latter errors appeared to be of an impulsive nature. The performance of the hyperactives also deteriorated more seriously over time than that of the normals. This deterioration seemed to be accompanied by increased motor restlessness which was measured by a stabilimetric cushion (Sprague and Toppe, 1966) attached to the child's chair.

Sykes also looked at the effect of piping intermittent white noise (80 decibels) into the experimental room at frequent random intervals while the subjects were performing on the continuous performance test. This attempt to distract the children did not affect the hyperactive subjects more than the normals. An attempt was also made to introduce conflicting cues on the choice reaction time test; in one series of trials the coloured background of the stimuli appearing on the screen was discrepant with the colour of the background of the target stimuli. This procedure also failed to affect the hyperactives differentially.

Nancy Cohen was also interested in attentional processes in children and her dissertation involved the monitoring of autonomic components of the orienting response and its habituation while the children were responding to different task demands (Cohen, 1970; Cohen and Douglas, 1972). She took skin response and heart rate measures under a variety of conditions: first during a rest period, then while the child was listening to stimuli to which he was not required to respond and, finally, while he was reacting to warning and reaction signals while performing on a delayed reaction time task.

Let me give you a brief summary of her results. First, there were no differences in the tonic levels of skin conductance or of heart rate during relaxation periods. Secondly, when the Ss were simply required to sit and listen to a series of tones which were presented through earphones there were no significant differences between hyperactive and normal subjects either on tonic levels of skin conductance or on the orienting response to the first or later tones. I would like to mention, however, that Cohen has some preliminary data which suggest that there may be differences when the tones are piped into the room through a loud speaker rather than through earphones. We plan to look more carefully into these effects.

When we look at the orienting response (OR) measures during the Delayed Reaction Time Task, where the Ss were required to make active responses to discrete stimuli, significant differences between the hyperactive and normal subjects emerge. The controls exhibited a significant increase in both tonic and phasic OR measures while the hyperactives remained relatively unresponsive. Performance on the reaction time task itself was also clearly deficient in the hyperactive children; compared with controls they exhibited slower reaction times and a greater amount of variability in performance. It would appear as if, for the hyperactive children, the warning signal given at the onset of each reaction time trial did not have the intended effect of alerting the child and preparing him to respond to the reaction signal.

It is important, however, to emphasize the extreme variability in the performance of the hyperactive children on this task. If we look at their

best trials, we find that they are capable of reacting as quickly as the normals; it is the erratic nature of their performance that reduces their scores. It is probably this kind of behaviour which inspires the lament so frequently heard from teachers: "He can do it if he wants to."

Cohen also attempted to improve the hyperactives' performance by reinforcing them for performing well. She did this simply by saying "good" on every trial on which the child surpassed his own basal reaction time score. After 15 trials with reward, the child was told that the examiner would no longer say anything but that he should still try to respond as quickly as possible. The reinforcement produced faster reaction times in both hyperactives and controls. There was also a decrease in variability in performance in both groups. However, when reward was withdrawn, the reaction times of the hyperactives tended to slow more quickly than those of the controls. The autonomic records during the reward condition also show differences between the hyperactives and controls. Although tonic activity increased for both groups when reward was introduced and stayed at the increased level when it was withdrawn, the pattern for the OR to the warning signal was quite different in the two groups. For controls, OR frequency increased significantly when reward was given and when it was withdrawn the frequency was still significantly higher than in the initial non-reward period. For hyperactives, on the other hand, OR frequency remained relatively low and did not change with the altered reinforcement conditions. While the tonic response is thought to measure a general increase in alertness or non-specific activation, the phasic response is thought to render the subject more sensitive to specific incoming stimuli (Lynn, 1966). It is this more specific response that seems to have been lacking in the hyperactive group. We feel that these findings underline the importance of training hyperactive children to concentrate on the critical aspects of a learning situation: it is apparently not sufficient just to increase general motivation. Recent findings by Penny Parry have also emphasized this point. She has discovered that a high level of general, non-contingent praise can even lead to a deterioration in performance.

Cohen also recorded other irrelevant movements while the subjects were performing on her task. The hyperactive children showed a strong tendency to press and release the response button after the appropriate response had been made: they also exhibited more intense movements in the left hand simultaneous with the appropriate response to the reaction signal with the right hand. One might think that such movements would interfere with the performance of these children. However, during reward, the frequency of these responses increased along with the improved performance. Thus, there does not seem to be a simple negative relationship between efficiency in attending and the amount the child moves. This and other observations we have made on these children have led us to question the importance that some educators and investigators have attached to making them sit still in order to get them to attend.

Susan Campbell's dissertation (Campbell, 1969; Campbell, Douglas, & Morgenstern, 1971) concentrated on variables quite different from the ones already discussed. Here, interest was in the style or approach that hyperactive children typically employ in problem solving. There has been a growing body of literature in the field of developmental psychology that has revealed the pervasive influence of several different cognitive styles on the approach that individuals take to a variety of problem solving situations. Several of the styles are relatively unrelated to intellectual ability. We felt that an understanding of the cognitive styles of hyperactive children might shed some light on their learning and behavioural problems and might also suggest directions for training and educational planning. Campbell studied four of the cognitive styles but I shall concentrate on two: field dependence-independence and reflection-impulsivity.

Kagan and his associates (Kagan, Rosman, Day, Albert, & Phillips, 1964) have demonstrated that one factor contributing to differences in cognitive functioning is the child's cognitive tempo, that is, his habitual speed of decision-making in situations with high response uncertainty. The impulsive child makes his decisions too quickly and, as a result, is likely to make errors, while the performance of the reflective child is characterized by long latencies and few errors. Normal children scoring high on impulsivity measures have been found to make more errors of commission on a serial learning task and to have higher error scores on tests of inductive reasoning. They have also been observed to be more distractable, less attentive, and more physically active than their more reflective peers (Kagan, 1965; 1966; Kagan, Pearson, & Welch, 1966). To measure reflection-impulsivity, Campbell used Kagan's Matching Familiar Figures Test which consists of sets of pictures of common objects and animals. The child is shown a standard stimulus and six similar ones and is required to choose the one picture from among the six alternatives which is identical with the standard. This is not particularly easy to do – many of the possible choices are so close to being correct that the child must hold back his tendency to respond long enough to look over the pictures very carefully. The test is scored for latency to first response and number of errors.

Witkins' dimension of field dependence-independence (Witkin, Dyk, Faterson, Goodenough, & Karp, 1962) reflects individual differences in the ability to separate an item from the field in which it is embedded. This dimension is also conceptualized in terms of differences in the degree to which perception is global and diffuse or structured and analytic. The field-independent individual is better able to overcome a confusing,

embedding context when isolating figure from ground. Witkin and his associates have demonstrated that field independent boys are more emotionally independent and display better control over impulses than field dependent boys (Witkin et al., 1962). Field-independent boys are also more concerned with intellectual tasks. Campbell used the Children's Embedded Figures Test (Karp & Konstadt, 1963) as her measure of field-independence. It consists of two series of simple figures embedded in more complex designs; the score is based on the number of figures correctly located.

Campbell found that both of the cognitive style measures differentiated significantly between hyperactives and controls. On the Matching Familiar Figures Test the hyperactives had significantly shorter latencies and made more errors. On the Embedded Figures Test they isolated fewer embedded figures. Thus, they would be classified as being more impulsive and more field-dependent than the normal controls. In her discussion of these results, Campbell suggests that it may be possible to modify the cognitive styles of hyperactive children. Training could be centred on teaching the child to delay responding until he has examined the various response alternatives in a situation. One could work up gradually from a few simple alternatives to several more complex ones. It might also be possible to use materials similar to the embedded figures so as to help him learn to focus on the essential features of a situation. He could be taught organized strategies of search for discovering the figure, first with simple examples, and later with more complex ones. Successful approaches for training children with attentional and impulsivity problems have been developed by Palkes, Stewart, and Kahana (1968), Santostefano and Stayton (1967), and, more recently, by Meichenbaum and Goodman (1971). We have also begun some preliminary work with a group of highly impulsive disadvantaged children. We have used games such as "May I," and a variety of card games and interesting projects in which the child can only succeed if he keeps his impulses under control. The teacher tries to teach and demonstrate "reflective" strategies and encourages the child to verbalize them and carry them out.

Before I leave Campbell's study, I want to mention just briefly that she also looked at a cognitive style called constricted-flexible control (Klein, 1954). She used the Colour Distraction Test developed by Santostefano and Paley (1964) which requires the child quickly to name the colour of objects arranged on a card, while ignoring distracting and contradictory stimuli. There was no evidence that the performance of the hyperactive children was impaired under either the distracting or interference conditions. Thus, this is one of several times we have failed to show any significant effect of distracting stimuli on our hyperactive children. We obviously need to learn a good deal more about attentional and distractibility problems in these children. I do want to emphasize, however, that their attentional problems appear even when they are working alone in a relatively empty, sound-proof room.

EFFECTS OF THE STIMULANT DRUGS

Now that we have established a pretty clear picture of some of the disabilities of hyperactive children, I would like to turn to a consideration of how drugs effect their performance on these same tasks. I am going to concentrate here on our studies using methylphenidate (Ritalin) which is one of the stimulants. We did do some earlier work with chlorpromazine (Freibergs, Douglas, & Weiss, 1968; Werry, Weiss, Douglas, & Martin, 1966), but found no evidence that it improved cognitive functioning. In the studies that I shall report the sample size varied from 20 to 50 subjects. Most of the studies used a double blind, cross-over design. The psychiatrist was allowed to control dosage to achieve a maximum effect. The dosage of methylphenidate was gradually increased from 5 mg/day to a maximum of 100 mg/day. Mean dosage was approximately 60 mg/day. Test-retest interval was two weeks.

The drug produced some positive effects on the tests in the clinical and learning disability battery: there was a significant increase in overall I.Q. and also on a few of the other tests in the battery; however, we could see no particular pattern in the kinds of tests affected. There was also significant improvement on several of the attention measures used by Sykes (Sykes, 1969; Sykes et al., 1971; Sykes et al., 1972). On the Continuous Performance Test on which the hyperactives had shown their most serious impairment, the number of correct responses increased, the number of incorrect responses decreased, and there was considerably less deterioration in performance over time when the children were receiving the drug. On the Serial Reaction Time Task, there was improvement on both correct and incorrect response measures; on the Choice Reaction Time Task, reaction time decreased. Thus, the drug produced positive changes on the measures on which the hyperactives had been most deficient but it also produced significant improvement on some measures on which there had been no significant differences between hyperactives and controls. On Cohen's delayed reaction time task (Cohen, 1970; Cohen, Douglas, & Morgenstern, 1971) the mean reaction times were faster and less variable and task-irrelevant motor responses were less frequent when the hyperactives were receiving the drug. On Campbell's cognitive style measures (Campbell, 1969; Campbell et al., 1971) the hyperactives became significantly less impulsive while on methylphenidate. This was reflected in

longer reaction times and fewer errors on the Matching Familiar Figures Test. The differences on the Embedded Figures Test of Field Dependence-Independence did not reach significance. Taken together, these results suggest to us that methylphenidate exerts its main effect by helping the hyperactive child sustain attention and control impulsivity. Similar results are being reported from several other laboratories (e.g., Conners & Rothschild, 1968; Knights & Hinton, 1969; Sprague, Barnes, & Werry, 1970).

SPECULATIONS ABOUT AETIOLOGY

There has been a good deal of physiological speculation about how the stimulants produce their apparently "paradoxical" effect and these theories, in turn, are tied to further speculations about the aetiology of the hyperactive disorder. Many clinicians and investigators believe that the hyperactive child has undergone some kind of brain damage; birth injuries appear to be the most frequently cited explanation (Laufer & Denhoff, 1957; Martin, 1967; Gross & Wilson, 1964; Rosenfeld & Bradley, 1948; Levy, 1959). It is, however, commonly recognized that no evidence of brain damage can be found in many of these children and so the more ambiguous notion of "brain dysfunction" has become popular (Clements & Peters, 1962; Stevens, Boydstun, Dykman, Peters, & Sinton, 1967). Recently, several writers have suggested the possibility of a biochemical defect (Shetty, 1971; Stewart, 1970; Silver, 1971); norepinephrine is the chemical most mentioned, partly because its release is thought to be stimulated by the amphetamines (Carr & Moore, 1969; Stein & Wise, 1969). Several of the investigators who have adopted a biochemical explanation believe that hyperactivity is a hereditary trait (Stewart, 1970; Silver, 1971). The fact that it is so much more common in males than in females suggests the possibility that it is a sex-linked or sex-influenced character; males, however, are also known to be more vulnerable to brain injuries.

There has been a good deal of theorizing about the nature of the brain mechanisms underlying the disorder. Among the possibilities that have been mentioned are the failure of some essential inhibitory control or filtering mechanism (Laufer, et al., 1957) and an imbalance between cortical and subcortical structures which results in the cortex having insufficient control over subcortical centres (Knobel, Wolman, & Mason, 1959). The reticular activating system plays an important role in several of these theories. Many of the theorists also think of the hyperactive child as suffering from either an abnormally high or abnormally low level of physiological arousal (Laufer, et al., 1957; Werry, Sprague, Weiss, & Minde, 1969; Stewart, 1970; Satterfield & Dawson, 1971). These beliefs are based on the children's high level of behavioural activity, the fact that the appear to be searching for sensory and kinesthetic input and the effectiveness of the stimulant drugs.

Our own research has added little that is definitive to these speculations. Members of our team have failed to find substantial evidence of brain damage in our children's birth histories, electroencephalograms or neurological examinations (Werry, Weiss, & Douglas, 1964; Werry, Weiss, Dogan, Minde, & Douglas, 1969; Minde, Webb, & Sykes, 1968), although, like the psychological tests, the neurological examinations have shown them to be poorly co-ordinated. It is important to remember, as I mentioned earlier, that we excluded from our samples children with clear evidence of brain damage: however no initial screening was done for more subtle signs. These studies have clearly emphasized to us the importance of proper control groups and follow-up studies in work of this kind; a surprisingly large number of our controls for example, had "abnormal" EEGs. We were also struck by the poor reliability of the electroencephalographic data over time. It should be noted, however, that we have not worked with the more sophisticated EEG techniques such as evoked potentials and contingent negative variation.

Although I am a little reluctant to add further to the rather loose physiologizing that has been going on, I would like to mention that Penny Parry has been experimenting with two tests which have been found to differentiate patients who have undergone frontal lobe surgery from patients who have lesions in other areas (Milner, 1963; Milner, 1964; Porteus, 1959). These are the Wisconsin Card Sorting Test which requires the ability to switch set and the Porteus Mazes which measure the ability to plan and follow rules. The hyperactives are doing very poorly on both of these tests, though at this point we are not sure that the reasons are always the same as those put forward by Milner (1964) to explain the poor performance of her frontal-lobe patients. I can only say at this time it is our impression that attentional and impulsivity problems contribute greatly to their performance. So far as "physiological arousal" is concerned, I am well aware of the complexities involved in using this term. However, for what it is worth, let me remind you that Cohen found no differences between resting levels of hyperactives and normals on either skin conductance or heart rate measures. Differences between the two groups showed up only on orienting response measures when the children were required to attend and respond to stimuli. However, this matter is far from settled. Although the findings of some other investigators (Stevens, Boydstun, Ackerman, & Dykman, 1968; Boydstun, Ackerman, Stevens, Clements, Peters, & Dykman, 1968) have tended to agree with our results, Satterfield & Dawson (1971) have reported first abnormally low and more recently (personal communication) abnormally high arousal levels in their hyperactive

groups. They have suggested that their conflicting findings may be explained by differences in the laboratory conditions in their two studies. As I mentioned earlier, physiological recordings in these children seem to be unusually sensitive to relatively minor changes in methodology. I suspect that this is due to the attentional problems I have described and also, perhaps, to a high degree of lability in both their behavioural and physiological responsivity. Their overreaction to praise which I mentioned earlier may be one example of this.

WHEN TO USE THE STIMULANTS?

It must be clear by now that although the physiological speculations about the hyperactive child raise many intriguing questions, they cannot offer any definitive answers. So far as the stimulant drugs are concerned, they seem to "work" with at least some of these children but we do not know why. This leaves the clinician open to severe ethical conflict about when the drugs should be used. Unfortunately, the question is in danger of becoming a political issue. I recently found myself embroiled in a debate with John Holt on this topic in the *New York Review of Books* (13 August, 22 October, and 3 December 1970). Not surprisingly, Holt places a good deal of blame for the problems of hyperactive children on the shoulders of unimaginative middle-class teachers; he also accuses the physicians who prescribe the drugs of "fashionable quackery." As one reads what he has to say, it becomes clear that no dull, middle-class professor ever taught him to do a literature review before allowing him to pontificate on a subject. Nevertheless, Holt and others who have voiced concern do raise a valid point. Although many of our children have expressed relief at finally being able to control themselves, I and most of my colleagues have come to believe that these drugs should be used only if the child's symptoms are extremely debilitating. Some of the stimulants are known to be addictive in adults and, though no cases of addiction have been reported by clinicians working with hyperactive children or adolescents, the possibility cannot yet be dismissed. We have also found greatly increased heart rates in some of the children in our short-term studies (Cohen & Douglas, 1972); we have no data as yet on long term effects. We find, too, that a few of our children become extremely depressed and show strangely flattened affect while they are on the drugs. Thus, we believe that the drugs should be used only after a very careful evaluation of the child's problems and if they are used, the physician should stay in close contact in order to titrate dosage and to monitor the youngster's response. It is to be hoped that, as our diagnostic techniques are sharpened, psychologists will be able to provide more help both in identifying children who may benefit from the stimulants and in evaluating their effects. As you might guess, I place most faith in methods for diagnosing attentional and impulsivity problems. However, though I believe this contribution would be a valuable one, my hope is that we will also be able to contribute to the development of training techniques which may at least diminish the need for drugs.

DOES THE HYPERACTIVE CHILD OUTGROW HIS SYMPTOMS?

I would like to look now at what we have been able to learn about the developmental history of the hyperactive disorder. As I have said, in the children we have studied, the symptoms have been present since infancy or early childhood. But what about the progress of the disorder as the child matures? There have been many clinical reports suggesting that these children "outgrow" their symptoms. If this were so, it might provide some evidence for the theories which suggest that the symptoms are a result of a "maturational lag" in the development of some critical part of the central nervous system (Lytton & Knobel, 1958; Solomons, 1965). Gabrielle Weiss has tried to stay in touch with our hyperactive children as they grow towards adolescence and we now have some data on a group of 13 to 16 year olds. There is some evidence, as there was in the school observations, that the hyperactivity has decreased; however, it is clear that impulsivity and inability to attend remain a problem. Cohen stayed on with the project after completing her degree and has run several of our tests in the longitudinal study (Cohen, Weiss, & Minde, in press). These young adolescents were still making more errors than their controls on the Continuous Performance Test, both on the visual and auditory form. On the Matching Familiar Figures Test of Reflection-Impulsivity they still reacted more quickly and they also made more errors than the controls on the Embedded Figures Test. Thus, they remained more impulsive and more field dependent. Cohen found no differences between normals and controls on the Stroop Colour-Word Interference Test. This, like Santostefano's Colour Distraction Test, which I mentioned earlier, is a measure of constricted-flexible control and is thought to measure the extent to which individuals are susceptible to cognitive interference from conflicting cues. This, then, represents another failure to prove that the hyperactives are unusually distractible. Psychiatric interviews revealed that 30 per cent of the children were described by their mothers as having no steady friends and 25 per cent had a history of fairly serious acting-out and anti-social behaviour (Weiss, Minde, Werry, Douglas, & Nemeth, 1971). Teachers' ratings and school reports analysed by Doba Lewin and Henry Lavigueur revealed that the children were still doing significantly worse than their controls at school, both academically and socially. Thus, these first follow-up investigations provide little hope that maturation is restoring these children to normality.

STOP, LOOK AND LISTEN!: AN UNDERLYING DIMENSION?

Let me try to summarize what we think we have learned thus far about the pattern of our hyperactive children's deficiencies. First, we should not forget that there are several areas of functioning that are relatively unimpaired in these children. Many of our subjects obtain average or even well above-average I.Q.s on standard intelligence tests, particularly if they are individually administered. As a group, they show no significant differences from normals in terms of language abilities, comprehension, or conceptual thinking. Neither is there any evidence of difficulty with short-term memory, even though we tested for this with several different kinds of materials. They appear, too, to be less disrupted by outside distractions than many of the reports in the literature would suggest; nevertheless they can be led astray by stimuli that are highly attractive to them. Although they do move more than other children, most of their behaviour seems to be directed to obvious goals, albeit their own. There is some suggestion, too, that activity level may not be the most critical aspect of the symptom picture of these children. For example, when they grow older they become less active but attentional and impulsivity problems remain.

As I looked back over our various studies, it struck me that one closely related group of characteristics can pretty well account for all of the deficiencies we have found. These youngsters are apparently unable to keep their own impulses under control in order to cope with situations in which care, concentrated attention, or organized planning are required. They tend to react with the first idea that occurs to them or to those aspects of a situation which are the most obvious or compelling. This appears to be the case whether the task requires that they work with visual or auditory stimuli and it also seems to be true in the visual-motor and kinaesthetic spheres. These same deficiencies – deficiencies which I have come to think of as the inability to "stop, look and listen" – seem also to influence the children's social behaviour. Penny Parry has shown that on a story completion task, they are unable to react realistically to a potentially frustrating situation, and, in real life, several of our older hyperactives are beginning to get into trouble with the law because of their inability to control their impulsive tendencies.

We have been struck by the degree to which our measures that tap attention, impulse control, and the ability to take an analytic approach to problems seem to go together in these children. These congruences were found in our early testing and the same constellation of deficiencies is present in our follow-up study. Furthermore, methylphenidate seems to exert its effects on these same measures. Since Sykes, Cohen, and Campbell all used the same subjects in their studies we were able to run correlations among their various measures; as expected, we found rather high intercorrelations. For example, correct responses on the continuous performance test correlated −0.66 with mean delayed reaction time and also −0.66 with errors on the Matching Familiar Figures Test.

RELEVANCE OF FINDINGS TO NORMAL POPULATION

All of this made us wonder whether scores from these same tests would be intercorrelated in a normal population. We were encouraged to think that this might be so by some findings which Campbell had obtained from a group of normal children (Campbell & Douglas, 1972). Her data had revealed significant correlations between her subjects' field dependence scores and their scores on the reflection-impulsivity measures. Furthermore, both measures were significantly correlated with scores on a story completion test that gauged the children's optimism about their ability to cope with a difficult, potentially frustrating situation. These correlations were independent of the effects of both intelligence and age. We therefore decided to investigate the possibility that the attention-impulse control constellation which seemed so important in the functioning of hyperactive children might also be an important factor underlying the cognitive and social functioning of normal children. We tested this out by administering to a group of 41 eight- and nine-year-old normal boys the same tests that we had found to differentiate between hyperactive and normal children. We also included in the test battery four I.Q. measures and predicted that these would be relatively unrelated to our hypothesized dimension. Finally, we added two anxiety measures which we also expected to be independent of our attention impulse control factor.

The results were extremely interesting. The correlational matrix revealed a very clear pattern of significant intercorrelations among most of the tests which we thought would tap our underlying "stop, look and listen" dimension. Very few of the measures correlated significantly with the I.Q. measures; there were also few significant correlations with the anxiety measures. A factor analysis was then performed; a varimax rotation produced four factors. The first of these looks very much like our hypothesized dimension. The tests loading significantly on it (in order of their loadings) included: the Porteus Mazes, the Children's Embedded Figures Test of Field Dependence-Independence, teachers' ratings on a hyperactivity scale, the eye-motor co-ordination subtest of the Frostig Motor Development Schedule, aggressive responses on a story completion test, the Bender

Visual-Motor Gestalt Test, a listening task involving loudness discrimination, the Matching Familiar Figures Test of Reflection-Impulsivity, the Continuous Performance Test, and responses to a story completion test which demonstrated the child's ability to cope realistically with a frustrating event. The second factor seems clearly to be an intelligence factor. Three of the intelligence tests loaded on it as well as two of the story completion test scores. The third is probably an anxiety factor. It loads on our two anxiety measures but also on the Lincoln-Oseretsky Schedule of Motor Development and the Bender Visual-Motor Gestalt Test.

These results have been obtained only recently and we have not had much chance to follow up their implications. At the moment we are inclined to think that our "Stop, Look and Listen" dimension is not unique to children diagnosed as hyperactive but is, rather, an important factor influencing the capacity of all children to cope effectively with a wide range of situations. We are interested in learning more about what the distributions of normal and hyperactive children look like on the various tests that define this factor. At the moment, I can only say that the distributions of scores of the "normal" sample on several of the tests seem to reveal considerable skewness on the "bad" end. It is as if something is responsible for producing more children who do badly on these tests than one would expect by chance.

Perhaps it may seem that I am pushing my "Stop, Look and Listen" dimension too far. However, before I conclude, I would like to try to stretch its implications to one more behavioural area – the area of moral behaviour. Michael Schleifer's doctoral dissertation (Schleifer, 1971; Schleifer & Douglas, in press; Schleifer & Douglas, unpublished manuscript) dealt with moral judgments in young children. His measures are based on Piaget's approach to moral development which parallels closely his theorizing about cognitive development. Piaget believes that the child's early moral judgments are limited by his cognitive capacities. For example, he tends to be very much influenced by the behaviour of authority figures because he sees moral rules as fixed laws. It is only later that he realizes that rules are made by parents who are not necessarily infallible. The young child is also limited by his "egocentrism." He is unable to take the viewpoint of another person and thus to take that person's intentions into account when he is making judgments about the goodness or badness of the person's acts. Piaget speaks also of "syncretism" a term which refers to the child's reacting globally to a situation rather than analysing its elements; the related concept of "centration" refers to the tendency to focus on some striking but superficial aspect of a phenomenon. By now it can be seen that we are approaching morality from the point of view of the ego rather than the superego. I hope that there can be detected in the above descriptions certain similarities with the kinds of abilities I have already discussed.

Schleifer took on the job of seeing whether there was, in fact, a relationship between our measures and the stages of moral development posited by Piaget. His test instruments were stories very similar to those used by Piaget and also, for his younger subjects, films which depicted situations similar to the ones in the stories. The stories and films portray situations in which a child commits an act such as breaking something. The subject is questioned about the goodness or badness of the child in the story and is asked to give the reasons behind his judgment. For example, there is one film in which a boy breaks several glasses while trying to help his mother with the dishes. In another film, a boy breaks a single glass while trying to steal cookies. Young children are likely to consider the first boy to be the really bad one because they fail to take intentions into account and also because they tend to be overly influenced by the amount of damage done. The children's responses to a series of stories (or films) were used to yield a "moral maturity score." The other measures included tests of reflection-impulsivity and field dependence-independence, as well as teachers' ratings on several behavioural scales.

Now for a brief summary of Schleifer's findings. In three different samples, significant correlations were obtained between moral judgment, field-dependence-independence and reflection-impulsivity scores. Most of the correlations ranged between 0.4 and 0.6. It is also important to note that none of the three measures correlated significantly with intelligence as measured by the Peabody Test of Verbal Intelligence. The data from teachers' ratings also lend support to our previous findings. In one sample, correlations between teachers' ratings of the child's propensity towards aggressive behaviour and the measures of moral judgment, field independence and reflectivity were negative and significant. In another sample, children high on the morality measure were also rated by their teachers as being more attentive and less impulsive.

Lest it is beginning to appear that I have got myself out on a limb, let me just say that I have some company. We recently discovered an article on "Conscience and Attentional Processes" by Grim, Kohlberg, and White (1968) in which the authors indulge in reasoning very similar to ours. They report on the chance meeting of two studies which happened to be going on in the same school. One group of investigators was studying attentional processes in children, using a reaction time task and associated galvanic skin response measures; their methods and measures were very similar to the ones used by Cohen in her dissertation. The other group of investigators was studying cheating behaviour, using situational tests of cheating and teachers' ratings of the children's ability to resist temptation.

As we would have predicted, there were several significant correlations between the measures obtained from the two studies. In discussing their results, the authors suggest that children who can maintain stable attention are also able to resist the quick, effortless solution obtained by cheating. The authors themselves are sufficiently honest and morally mature, however, to admit that they are not the first psychologists to see the theoretical connection between attentional processes and moral behaviour. Perhaps you will recognize the author of the following quotation: "If a brief definition of ideal or moral action were required, none better would fit the appearance than this: it is action in the line of greatest resistance ... We reach the heart of volition when we ask by what process it is that the thought of any given object comes to prevail stable in the mind. Attention with effort is all that any case of volition implies. The essential achievement of will is to *attend* to a difficult object and hold it fast before the mind."

I am sure you will have guessed that the author is William James (1890). So, like many psychologists pushing their theories, I have discovered that James got there before me!

RÉSUMÉ

Exposé d'un programme de recherche pour étudier les déficits spécifiques des enfants hyperactifs et l'efficacité des agents psychoanaleptiques sur les mesures prises en vue de distinguer normaux et hyperactifs. Il est allégué qu'un noyau de symptômes centrés sur l'incapacité de maintenir l'attention nécessaire à contrôler ses impulsions expliqueraient la plupart des déficits observés dans le groupe des hyperactifs. Il semble également que ce soit sur ces déficits que les stimulants exercent leur effet principal. Des études corrélationnelles et factorielles suggèrent que la même constellation d habiletés imprègne le comportement des enfants normaux dans plusieurs domaines du fonctionnement cognitif et social.

REFERENCES

BOYDSTUN, J.A., ACKERMAN, P.T., STEVENS, D.A., CLEMENTS, S.D., PETERS, J.E., & DYKMAN, R.A. Physiologic and motor conditioning and generalization in children with minimal brain dysfunction. *Conditional Reflex*, 1968, 3, 81–101.

BURKS, H.F. The hyperkinetic child. *Exceptional Children*, 1960, 27, 18–20.

CAMPBELL, S. Cognitive styles in normal and hyperactive children. Unpublished doctoral dissertation, McGill University, 1969.

CAMPBELL, S.B., & DOUGLAS, V.I. Cognitive styles and responses to the threat of frustration. *Canadian Journal of Behavioural Science*, 1972, 4, 30–42.

CAMPBELL, S.B., DOUGLAS, V.I., & MORGENSTERN, G. Cognitive styles in hyperactive children and the effect of methylphenidate. *Journal of Child Psychology and Psychiatry*, 1971, 12, 55–67.

CARR, L.A., & MOORE, K.E. Norepinephrine release from brain by d-amphetamine in *vivo*. *Science*, 1969, 164, 322–323.

CLEMENTS, S.D., & PETERS, J.E. Minimal brain dysfunctions in the school-age child. *Archives of General Psychiatry*, 1962, 6, 185–197.

Learning, reaction time, activity, and classroom behavior in disturbed children. *American Journal of Orthopsychiatry*, 1970, 40, 615.
SPRAGUE, R.L., & TOPPE, L.K. Relationship between activity level and delay of reinforcement in the retarded. *Journal of Experimental Child Psychology*, 1966, 3, 390–397.
STEIN, L., & WISE, C.D. Release of norepinephrine from hypothalamus and amygdala by rewarding medial forebrain bundle stimulation and amphetamine. *Journal of Comparative and Physiological Psychology*, 1969, 67, 189–198.
STEVENS, D., BOYDSTUN, J., ACKERMAN, P., & DYKMAN, R. Reaction time, impulsivity, and autonomic lability in children with minimal brain dysfunction. *Proceedings of the seventy-sixth annual convention of the American Psychological Association*, 1968, 3, 367–368.
STEVENS, D.A., BOYDSTUN, J.A., DYKMAN, R.A., PETERS, J.E., & SINTON, D.W. Presumed minimal brain dysfunction in children (Relationship to performance on selected behavioral tests). *Archives of General Psychiatry*, 1967, 16, 281–285.
STEWART, M.A. Hyperactive children. *Scientific American*, 1970, 222, 94–98.
STRAUSS, A.A., & LEHTINEN, L. *Psychopathology and education of the brain injured child.* New York: Grune & Stratton, 1947.
SYKES, D.H. Sustained attention in hyperactive children. Unpublished doctoral dissertation, McGill University, 1969.
SYKES, D.H., DOUGLAS, V.I., & MORGENSTERN, G. The effect of methylphenidate (Ritalin) on sustained attention in hyperactive children. *Psychopharmacologia*, 1972, 25, 262–274.
SYKES, D.H., DOUGLAS, V.I., WEISS, G., & MINDE, K. Attention in hyperactive children and the effect of methylphenidate (Ritalin). *Journal of Child Psychology and Child Psychiatry*, 1971, 12, 129–139.
WEISS, G., MINDE, K., WERRY, J.S., DOUGLAS, V.I., & NEMETH, E. Studies on the hyperactive child: A five year follow-up. *Archives of General Psychiatry*, 1971, 24, 409–414.
WERRY, J.S., & SPRAGUE, R.L. Hyperactivity. In C.G. Costello (Ed.), *Symptoms of psychopathology*. New York: Wiley, 1970.
WERRY, J.S., SPRAGUE, R.L., WEISS, G., & MINDE, K. Some clinical and laboratory studies of psychotropic drugs in children – an overview. In W.L. Smith (Ed.), *Symposium on higher cortical functions*. Springfield, Illinois: Thomas, 1969.
WERRY, J., WEISS, G., DOGAN, K., MINDE, K., & DOUGLAS, V.I. Studies on the hyperactive child. VII. Comparison of neurological findings between hyperactive, normal and neurotic children. Paper read at the Canadian Psychiatric Association Annual Meeting, Toronto, June 1969.
WERRY, J., WEISS, G., & DOUGLAS, V.I. Studies on the hyperactive child I: Some preliminary findings. *Canadian Psychiatric Association Journal*, 1964, 9, 120–130.
WERRY, J.S., WEISS, G., DOUGLAS, V.I., & MARTIN, J. Studies on the hyperactive child. III: The effect of chlorpromazine upon behaviour and learning ability. *Journal of American Academy of Child Psychiatry*, 1966, 5, 292–312.
WITKIN, H.A., DYK, R.B., FATERSON, H.F., GOODENOUGH, D.R., & KARP, S.A. *Psychological differentiation*. New York: Wiley, 1962.

First received 15 May 1972

Sex differences in personality structure at age 14*

PHILIP E. VERNON
University of Calgary

ABSTRACT

Though much work has been published on differences between the sexes in Western countries on tests of abilities, interests, and personality traits, little attention has been paid to the organization or structuring of psychological variables among males and females separately. An investigation of nearly 400 Grade 8 students in a Canadian city, including some 10 hours of testing, indicated that the ability factors in the two sexes are closely similar, though their relations to other variables such as age, artistic and scientific interests, social attitudes, and personality tendencies, often differ markedly. For example the psychological significance of field independence and of divergent thinking in the adolescent personality, and the organization of career and other interests, were so different that mixed-sex studies should be discouraged.

Most textbooks, or chapters of books, on individual differences devote considerable space to differences between the sexes in abilities, interests and personality traits (e.g. Anastasi, 1958; Tyler, 1965), and there is no need to recapitulate the well-established facts. Many published tests, at least for adolescents or adults, also provide separate norms for the sexes. Yet it is generally assumed that the main dimensions or factors of ability or personality have the same significance in males and females, and are similarly organized and structured. For example the Primary Mental Abilities batteries, the MMPI, the Study of Values, and Cattell's personality questionnaires are presumed to measure the same abilities or traits in each sex. Counselling or clinical psychologists would, however, be more likely to deny that anxiety, say, or interest in a scientific career, do mean the same thing psychologically in a girl as in a boy, or in a woman as in a man. Likewise, the developmental psychologist tends to find different influences operating in the personality development of boys and girls (cf. Honzik, 1967), and much has been written on the differing identification processes in the acquisition of male and female sex roles (e.g. Kagan, 1964). Rather generally it seems possible to trace fairly straightforward connections between, say, maternal warmth or rejection and the personal characteristics of boys; whereas in girls, personal behaviour depends more on the immediate social context (cf. Schaefer & Bayley, 1963). Different personality patterns, again, are associated with verbal, number, and spatial abilities in boys and girls, according to Ferguson and Maccoby (1966).

* The writer is greatly indebted to The Canada Council, Social Sciences Division, for financial support of this study. A brief version of this paper was presented at the Canadian Psychological Association meeting in Montreal, June 1972.

825

final

Impulsivity - reflectivity and differential reinforcement of low rates (DRL) performance

Norman Stein and Richard Landis

Syracuse University

Paper presented at the meeting of the Eastern Psychological Association, Washington, D.C., 1973.

54

The present study is derived from two, largely independent, areas of research; first, the cognitive style literature on impulsivity--reflectivity and, second, the operant conditioning literature on differential reinforcement of low rates (DRL) schedules.

Regarding the area of cognitive style, Kagan (1965) has isolated a conceptual tempo dimension called "impulsivity--reflectivity" which underlies modes of problem solving in children. The task used to differentiate these groups is called the Matching Familiar Figures (MFF) Test (See Figure 1). On this test the child points out the one stimulus, from among six highly similar variants, that is identical to a standard. Subjects scoring above the sample median for number of errors and below the median for decision time are designated as impulsive, while subjects scoring above the median for decision time and below the median for number of errors are designated as reflective. Research has consistently shown that impulsive children perform relatively poorer across a variety of motor, perceptual, and conceptual tasks which typically require an inhibition of the choice response in order to optimize performance. If this is the case, then the conceptual tempo dimension should be predictive of differences in performance on a learning task which requires temporal spacing of responses in order to maximize reinforcement.

One such task is derived from operant conditioning methodology and consists of responding on a telegraph key programmed to provide reinforcement according to a DRL 6 sec. schedule. The schedule provides reinforcement contingent upon responses which terminate inter-response times (IRTs) of 6 sec. or longer. Responses terminating IRTs shorter than 6 sec. are not reinforced and, in addition, reset the timer governing reinforcement.

Thus, in order to maximize reinforcement, a child must refrain from engaging in the reinforced response for at least 6 sec. The alternate behavior which occurs during the IRT is called collateral behavior. In the present experiment, collateral behavior is measured by presses on either of three telegraph keys which are adjacent to the directly reinforced response - key. Previous research has shown that under such an arrangement, accomodation to the DRL schedule requirement is accompanied by regular and frequent occurrences of unscheduled collateral behavior and that this collateral behavior, through the effects of adventitious reinforcement, can serve a mediating function that regulates the temporally discriminated performance (Stein and Landis, 1973).

The present study documents the performance of cognitively impulsive and reflective children operating under a DRL schedule of reinforcement by measuring response rates, reinforcement rates, reinforcements per response rate (i.e., efficiency), distribution of IRTs, and the behavioral mechanisms used to mediate successively reinforced responses. In addition, it sought to examine the role of instructions about the requirements for reinforcement on these variables.

METHOD

Subjects: The MFF was administered to 73 third-grade boys of a suburban elementary school. From this pool, 20 impulsive and 20 reflective children were identified. The mean age was 8.94 years and the mean IQ was 107.6. There were no significant differences between reflective and impulsive Ss on these characteristics.

Apparatus. The Ss were seated before a chassis containing four adjacent telegraph keys mounted 9.5 cm apart. Each key required a downward force of 438 gm to be electronically recorded as a response. A four-digit add counter was mounted in the middle of the chassis to indicate the cumulative number of reinforcements (points). Points were backed up by M & M candy reinforcers. White noise (60 db) was continuously presented through headphones worn by the S and a sound attenuating chamber housing the programming equipment acoustically isolated the S from the equipment. The third key from the Ss' left (Key 3) was programmed to produce reinforcement according to a DRL 6-sec. schedule.

Procedure. The general design involved four equal sized groups: two groups of impulsive Ss, and two groups of reflective Ss, one group from each received instructions about the DRL requirements and one group received no instructions about these requirements. Initially, to all Ss, the task was introduced as a game in which they should try to cumulate points and that each point equalled one M & M. Each point was also to be accompanied by a flash of a red light above the counter. Next, to assess S's baseline rate of key pressing all Ss were asked to practice using the keys but told they would not be able to obtain points at this time. After a two minute baseline phase, the two Instruction Groups were explicitly told the nature of the contingencies that would be in effect during the next 30 minutes. Essentially, they were told "...if you press the second key from your right, wait for 6 seconds before you press it again, then you'll get a point every time. If you press it before 6 seconds are up, then you will not get a point and will have to wait another 6 seconds before you can get a point."

RESULTS

Figure 2 shows the mean number of responses on Key 3, the key programmed on the DRL 6 sec. schedule, during baseline and during successive 5 minute intervals throughout training. No difference in response rate between the groups occurred during baseline ($F = 0.31$, $df = 3/36$, $p > .10$). However, throughout DRL training significant main effects were obtained for Instructions ($F = 6.65$, $df = 1/36$, $p = < .01$), Conceptual Tempo ($F = 4.51$, $df = 1/36$, $p = < .04$), and Time ($F = 3.19$, $df = 5/180$, $p = < .01$). Comparison of group means by t-tests showed that instructions about the contingencies significantly reduced the response rate of impulsives (Group II) compared to uninstructed impulsives (Group INI) ($p < .005$). The Uninstructed Reflectives (Group RNI) did not differ significantly from the Instructed Reflectives (Group RI). Though both instructed groups responded with similar rates, the difference between INI and RNI groups was significant ($p < .01$). It was also found that while the rate of responding decreased from the first to the last 5 minutes of training for both the INI ($t = 2.73$, $df = 180$, $p = < .01$) and RNI ($t = 2.50$, $df = 180$, $p = < .05$) groups, both instructed groups responded at essentially the same low rate throughout DRL training.

Figure 3 shows the mean number of collateral responses per minute on the three nonreinforced keys during baseline and during successive five minute blocks throughout DRL training. Again, no significant difference between groups was obtained during the baseline period ($F = 0.59$, $df = 3/36$, $p > .10$). During DRL training, there was a significant main effect for Instructions ($F = 13.04$, $df = 1/36$, $p = < .001$), in which the RNI Group responded with reliably higher rates than the II Group ($p < .005$). Similarly, the RNI Group emitted significantly more collateral responses than the RI group ($p < .05$). Significant interactions were also obtained for both Time X Instructions ($F = 4.14$, $df = 5/180$, $p < .001$) and Time X Conceptual Tempo ($F = 2.66$, $df = 5/180$, $p < .02$). Analyses of simple main effects revealed that while the two instructed groups responded with significantly lower rates than the two uninstructed groups during each time period ($p < .05$), only the uninstructed groups showed significant changes over time ($F = 16.36$, $df = 5/180$, $p < .01$). Thus, together with the findings on Key 3 responding, it was found that only the uninstructed groups exhibited changes in the rate of responding during exposure to the DRL contingencies. However, while the collateral response rates significantly increased from the beginning to the end of training for both the INI ($t = 4.73$, $df = 180$, $p < .001$) and the RNI ($t = 7.36$, $df = 180$, $p < .001$) groups, their rate of Key 3 responding had significantly decreased during the same time periods.

The direct comparisons between the rate of responding on Key 3 and on the collateral keys (1, 2, 4) for all groups are seen clearly in Figure 4. Series of one way ANOVAs indicated no differences within the INI, II, and RI groups in the response rates across the keys. However, the RNI Group tended to respond less frequently on Key 3 than on the collateral keys ($F = 2.68$, $df = 3/36$, $p < .10$) and at a rate comparable to that of the instructed groups.

The proportion of IRTs ≥ 6 sec for Key 3 responses, which reflects the efficiency of performance — that is, the percent of Key 3 responses that were reinforced, is shown in Figure 5. The same essential and almost identical relationships, were obtained for the other number of reinforcements per minute. Again, the four groups were found to be comparable in the proportion of IRTs ≥ 6 sec emitted during baseline ($F = 1.45$, $df = 3/36$, $p > .10$). A significant main effect was obtained for Instructions for both the efficiency

($F = 17.44$, $df = 1/36$, $p = < .001$) and the reinforcements per min ($F = 10.19$, $df = 1/36$, $p = < .003$) measures. Instructions significantly increased the efficiency ($p < .02$) and the number of reinforcements obtained ($p < .005$) by Impulsives. In contrast, Instructions enhanced the efficiency by Reflectives ($p < .03$) but did not correspondingly increase the rate of reinforcement ($p > .15$). The Time x Instructions interaction was also significant for the efficiency ($F = 3.77$, $df = 5/180$, $p = .003$) and reinforcements per min ($F = 4.21$, $df = 5/180$, $p = .001$) measures. Comparisons of the first and last time blocks revealed that both the INI and RNI Groups became more efficient and earned reinforcement at a faster rate as a result of extended exposure to the DRL contingency ($p < .001$), the RI Group also improved significantly on these measures ($p < .01$), while the II Group responded no more efficiently and earned no more reinforcements at the end than at the beginning of DRL training.

Although the INI group exhibited a significantly higher Key 3 response rate than the RNI group, the main effects for conceptual tempo on the efficiency ($F = 3.04$, $df = 1/36$, $p = .08$) and reinforcement rate ($F = 3.03$, $df = 1/36$, $p = .08$) measures failed to attain traditionally acceptable level of significance. Nevertheless, further analyses revealed that while the Instructed groups did not differ, the RNI Group responded more efficiently ($p < .05$) and obtained more reinforcements ($p < .025$) than the INI Group.

To assess the structure of the temporal performance for each group, the relative frequencies of IRT and the IRT/CP on Key 3 were obtained for the baseline period and for the first, middle, and last five min blocks of training. The results are shown in Figure 6. As shown, during Baseline, Impulsives emitted relatively fewer long IRTs than Reflectives and, given the opportunity, had a high probability of terminating IRTs of about 4 sec. The IRT/OP distributions for the Reflectives during Baseline were flat, indicating that the probability of emitting an IRT of any duration between 0-16 sec was the same. Most noticeable was the immediate elimination of short IRTs by the introduction of instructions and the rapid development of an accurate temporal discrimination by the Instructed Reflective and Impulsive Groups. The IRT distributions of the II group remained essentially unchanged throughout training while the RI Group showed a more gradual development of a precise discrimination. In contrast, the Uninstructed Impulsive and Reflective Groups generated considerably more IRTs in the shorter categories throughout training. In neither group was bursting (i.e., 0-2 sec IRTs) eliminated entirely; however, for Group INI, approximately 30% of all IRTs were in the 0-2 sec category during the last five min of training. In contrast, this percentage is equalled by the RNI Group only during the first five min of training. Thereafter, the percentage of 0-2 sec IRTs decreased steadily until, at the end of training, only 10% of the IRTs were less than 2 sec.

The proportion of bursts (i.e., a sequence of two or more responses with IRT $\leq$ 2 sec.) immediately following reinforced responses was determined. It was found that, on the average, for Group INI 31% of the reinforced responses were followed immediately by bursting while for Groups II, RNI, and RI the mean proportions were 2%, 6% and 0%, respectively. Statistical analysis indicated no differences between these latter groups but each was significantly different from Group INI.

DISCUSSION

Since there were no differences between reflective and impulsive Ss on the baseline measures of motor activity, the differences in responding between these groups during DRL must be accounted for in terms other than simply differential predispositions to respond at faster or slower rates. The present results suggest that whether or not cognitively impulsive and reflective children exhibit high or low response rates depends upon the "stimulus structure" or lack of ambiguity present for the child in the learning situation. When the DRL task was structured and made unambiguous through explicit instructions about the requirements for reinforcement, there were no reliable differences between impulsives and reflectives. However, when the DRL task was unstructured and ambiguous, such that the relationships between the task, the S's behavior, and the consequences of the behavior were unspecified, reliable differences between reflective and impulsive children emerged. Under these latter conditions, impulsive children emitted higher rates of the reinforced response, tended to obtain fewer reinforcements, were less efficient, and responded with equal frequencies on all the response keys. In contrast, uninstructed reflectives emitted a significantly lower response rate on the reinforced key in comparison to the collateral keys.

The higher response rate exhibited by the uninstructed impulsives was not a function of a higher general arousal or activity level. Rather, the higher rates were the result of rapid repetitions of the reinforced response (i.e., response bursts) and their occurrence was specifically related to the effects of reinforcement. Rapid bursts of responding typically occurred under two conditions: (1) for both uninstructed reflective and impulsive children bursting

on S^D for responding

62

$S^R \cdot S^P \rightarrow R$

occurred following nonreinforced Key 3 responses, and (2) for uninstructed impulsive children only, bursting also occurred following reinforced Key 3 responses. In the first instance, bursting probably reflected an accommodation to an unreceived but anticipated reward. In the second instance, which was specific to uninstructed impulsive Ss, bursting appeared as an operant under the discriminative control of the prior reinforcement. Thus, for uninstructed impulsive Ss, reinforcement may have served as an S_D for rapid repetition of the same response. In contrast, with uninstructed reflectives, reinforcement may have served as an S_D to engage in collateral behavior which effectively mediated the IRT.

The observation that reinforcement may control different behavioral tendencies for reflective and impulsive children in previously unstructured situations is also relevant to the proposal that impulsives fail to discriminate as accurately those situations in which it is appropriate to respond slowly. The data shown in Fig. 4 support this hypothesis, such that the INI Group responded with equal frequency on all keys, while the RNI Group responded less frequently on the reinforced key than on the collateral keys. Thus, impulsives exhibited a flat or broad gradient of generalization and reflectives exhibited a sharp gradient of generalization with a peak at the key directly associated with reinforcement. Since a steep generalization gradient reveals the presence of a discrimination, it is possible that reflectives discriminated stimuli associated with reinforcement from those associated with non-reinforcement. The broad generalization gradient obtained from the uninstructed impulsive Ss reveals the absence of such a discrimination. This interpretation, however, is limited by the finding that impulsives obtained fewer reinforcements than

FIG. 2

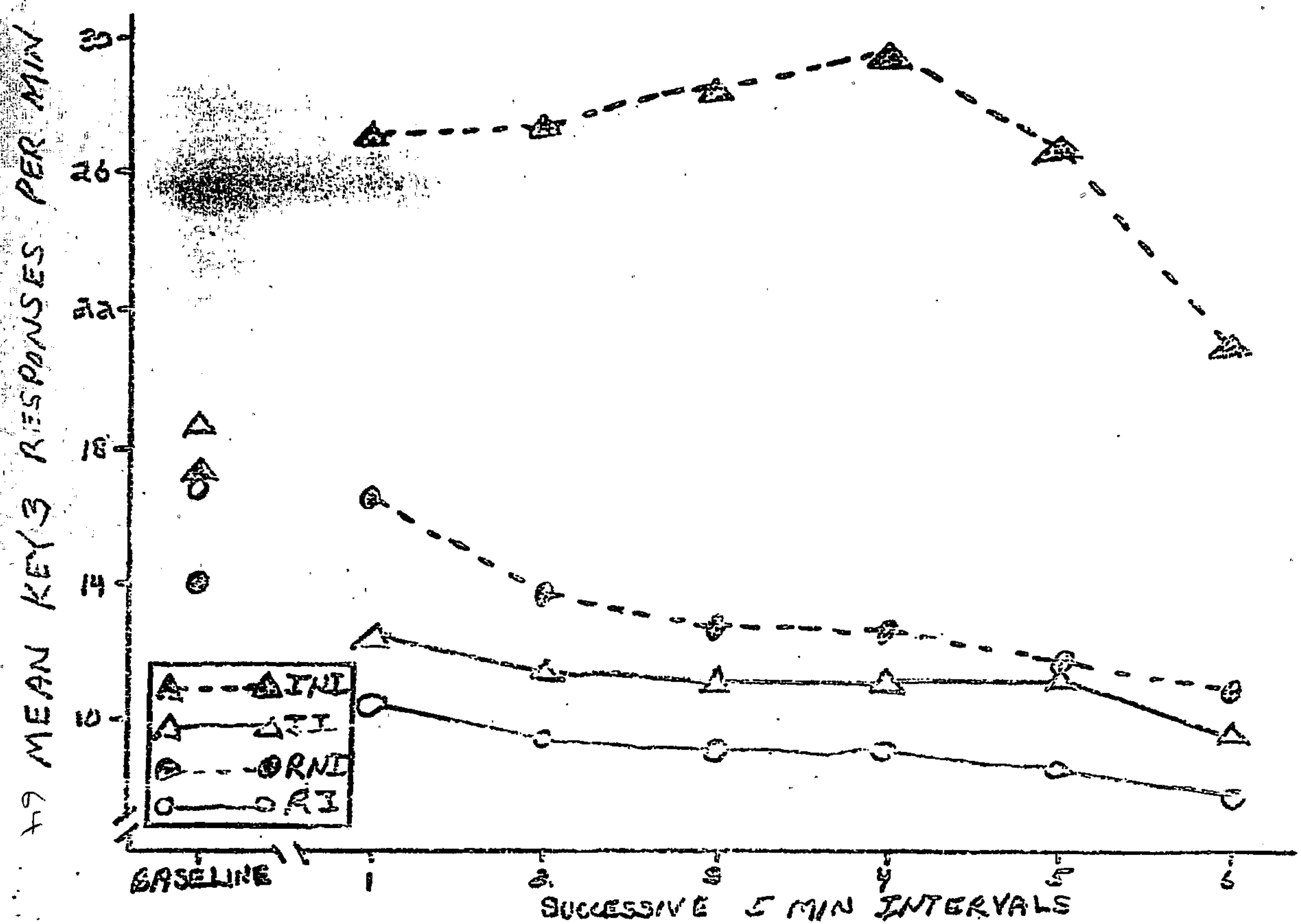

MEAN KEY 3 RESPONSES PER MIN
30
26
22
18
14
10
INI
RI
RNI
RI
BASELINE
1
2
3
4
5
6
SUCCESSIVE 5 MIN INTERVALS

FIG. 3

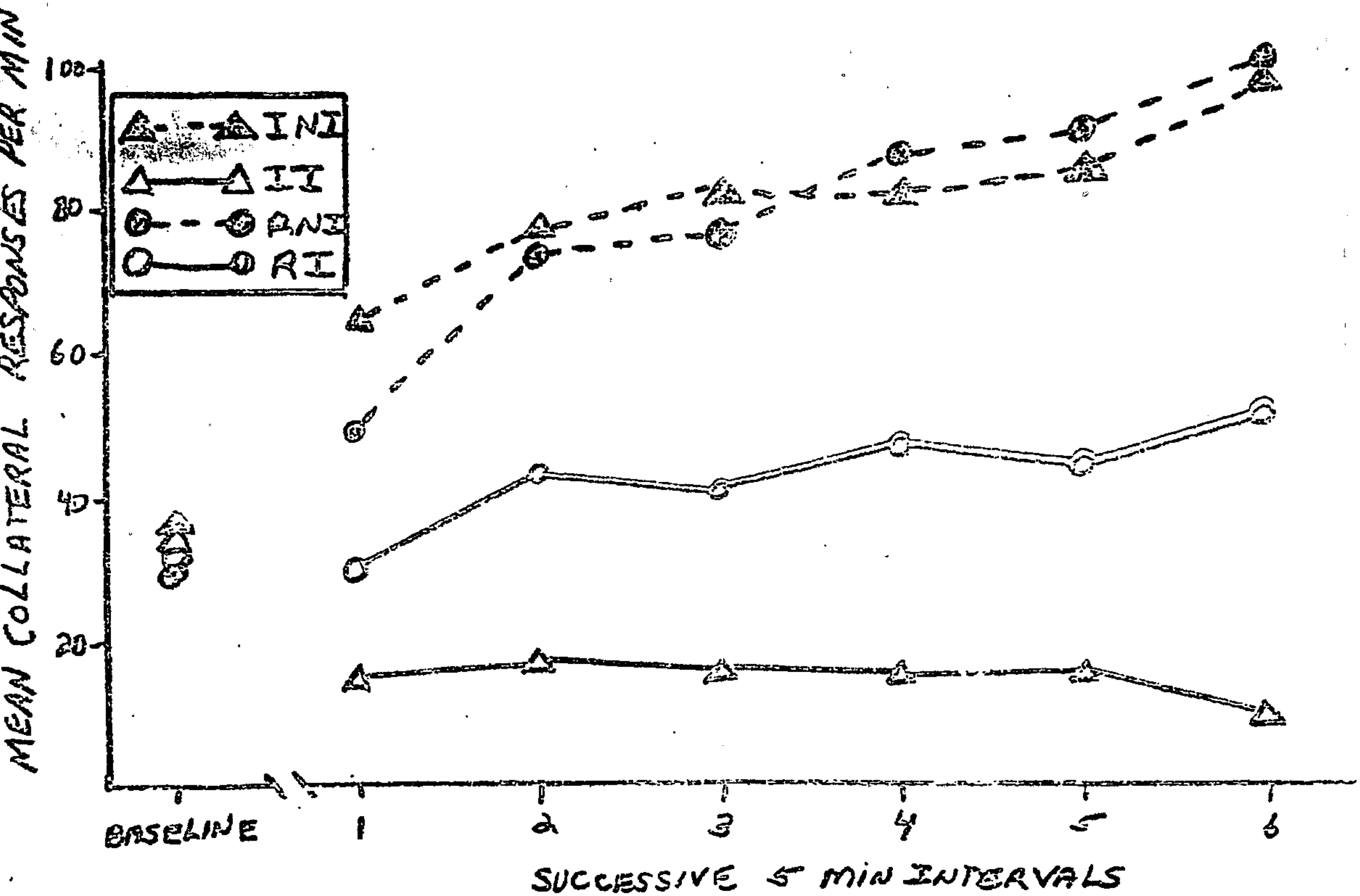

INI
II
ANI
RI
MEAN COLLATERAL RESPONSES PER MIN
100
80
60
40
20
BASELINE
1
2
3
4
5
6
SUCCESSIVE 5 MIN INTERVALS

FIG. 4

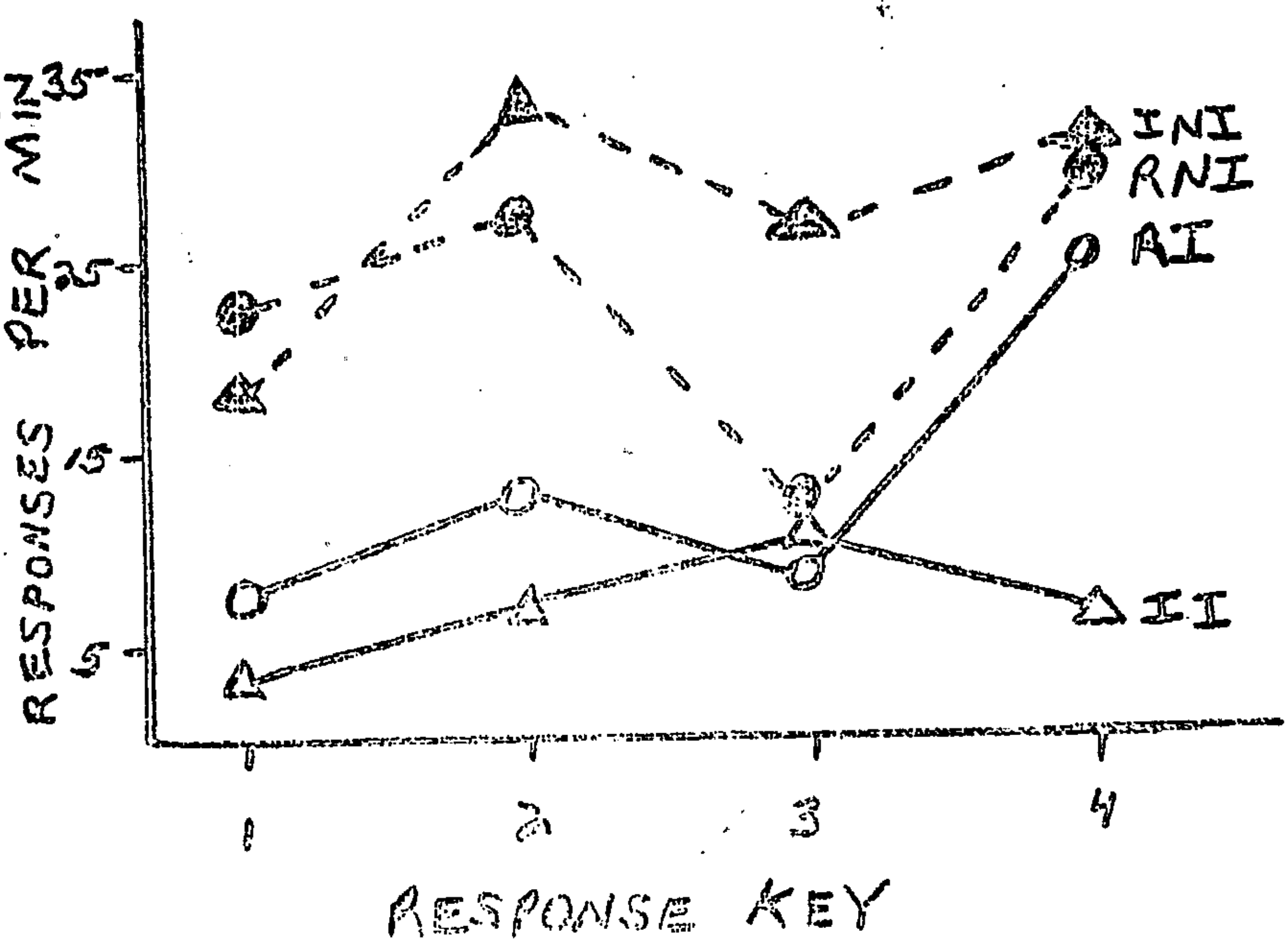
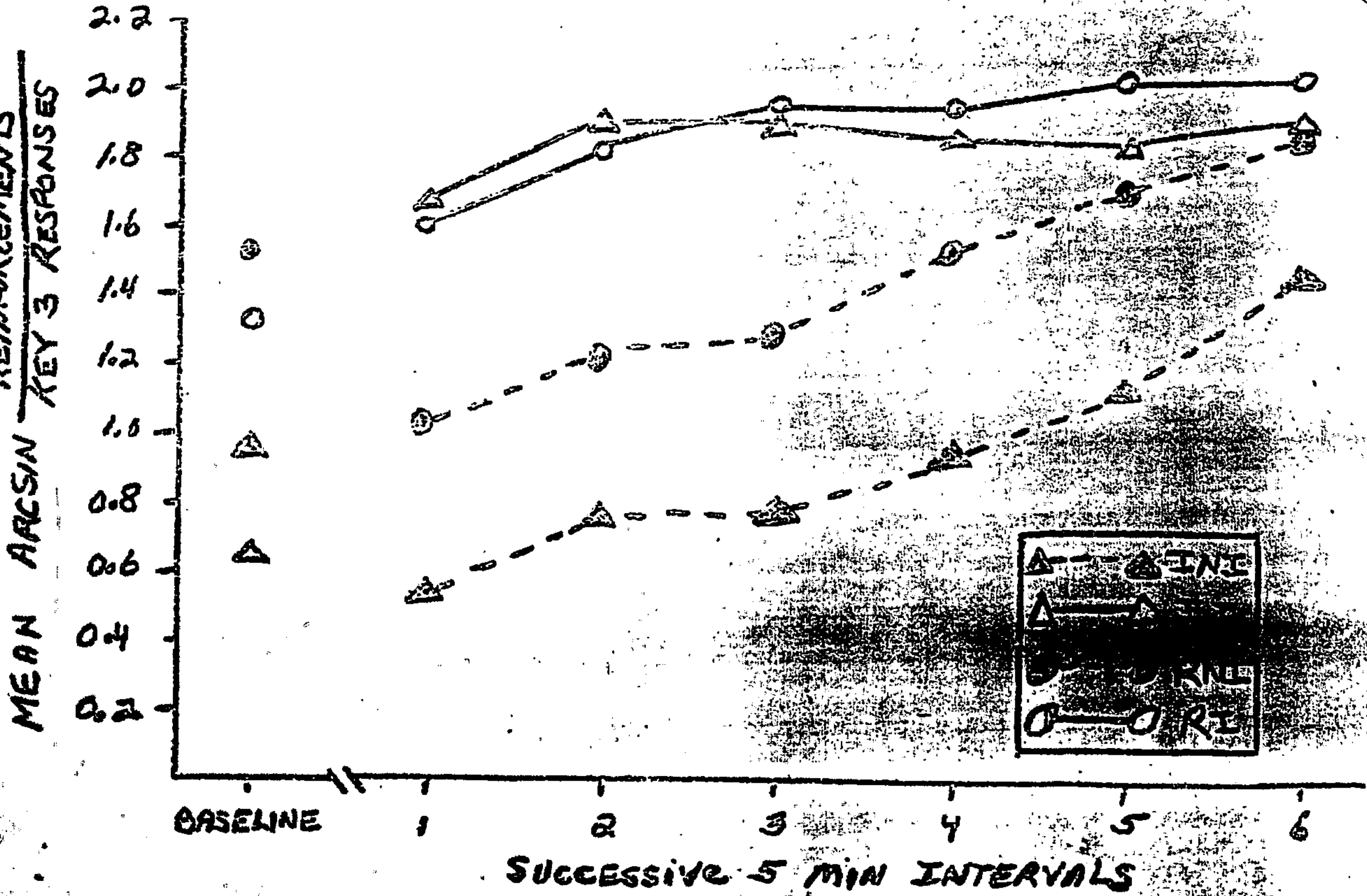


FIG. 5


MEAN ARCSIN REINFORCEMENTS / KEY 3 RESPONSES
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
BASELINE
1
2
3
4
5
6
SUCCESSIVE 5 MIN INTERVALS
INI
RI

reflectives. Thus, it is unclear whether impulsives obtained fewer reinforcements because they may fail to discriminate between stimuli associated with reinforcement from those which are not, or fail to discriminate because they receive fewer reinforcements.

REFERENCES

Kagan, J. Impulsive and reflective children: Significance of conceptual tempo. In J. Krumboltz (Ed.) Learning and the educational process. Chicago: Rand McNally, 1965.

Stein, N. and Landis, R. Mediating role of human collateral behavior during a spaced-responding schedule of reinforcement. Journal of experimental psychology, 1973, 97 (1), 28-33.

See page 131

851

J. Child Psychol. Psychiat., Vol. 12, 1971, pp. 129 to 139. Pergamon Press. Printed in Great Britain.

ATTENTION IN HYPERACTIVE CHILDREN AND THE EFFECT OF METHYLPHENIDATE (RITALIN)*

DONALD H. SYKES,† VIRGINIA I. DOUGLAS, GABRIELLE WEISS and KLAUS K. MINDE

Departments of Psychology and Psychiatry, McGill University and the Montreal Children's Hospital, Montreal, Canada

CHILDREN are frequently described as overactive, particularly by parents and teachers. Lapouse and Monk (1958) estimated that some 50 per cent of all children between the ages of 6 and 12 yr are described by their mothers as highly active although their activity is not necessarily a problem. Stewart *et al.* (1966) have estimated that 4 per cent of all school-age children show a level of activity sufficiently excessive and sustained as to be a serious source of complaint at both home and school. Such children have been labelled "hyperactive" by clinicians.

Given the problems surrounding the definition and measurement of activity level, particularly in humans (Cromwell, Baumeister and Hawkins, 1963), it is by no means surprising that there is little agreement on the definition of hyperactivity. For some investigators it implies a greater quantity of movement (Schulman, Kaspar and Throne, 1965), but this has been questioned by others (Werry and Sprague, 1970). Cromwell *et al.* (1963) have suggested that the "overactivity" of "hyperactive" children may be a reflection of the short attention span and rapidly changing goal directions of such children. Thus, these investigators argue that hyperactive children may be thought of as children whose behaviour is fragmented or disorganized and continually changing direction such that an impression of a high level of activity is created.

Certainly, the inability of hyperactive children to maintain their attention toa task has frequently been remarked in the clinical literature. However, other than the clinical statements and the remarks of teachers and parents, there are few experimental studies which clearly show an attentional deficit in such children. In fact, there is no reliable evidence that they are impaired on the usual tests of attention, such as the coding subtest of the Wechsler Intelligence Scale for Children (Douglas, Weiss and Minde, 1969). It has been suggested with respect to this latter test, that if the attentional impairment took the form of brief lapses in attention, then performance on the coding test need not necessarily be impaired. The nature of the

*The research was supported by a research grant from the Ciba Pharmaceutical Company, Dorval, Quebec, to Dr. G. Weiss.

†Present address: Department of Psychology, The Queen's University of Belfast, Belfast, Northern Ireland.

Accepted manuscript received 16 *April* 1971

69

coding test is such as to allow the subject to compensate for any brief lapses by working rapidly in between lapses (Rosvold *et al.*, 1956). In contrast, impaired performance due to momentary lapses in attention should be noted on a task in which the arrival of significant information is unpredictable. A task designed to meet these requirements was constructed by Rosvold *et al.* (1956), namely the Continuous Performance Test.

This task requires the subject to monitor a screen on which letters appear at regular intervals and to make a response whenever a previously specified stimulus appears. The test is similar to many of the tasks used in vigilance studies. Any momentary lapse in attention which is coincident with the appearance of a significant stimulus would appear as an error of omission. Thus, the primary purpose of the present study was to examine the performance of a group of hyperactive children relative to the performance of a group of normal control children on a task susceptible to momentary lapses in attention.

It has been observed that motor restlessness in adults increases with time on a vigilance task (Baker, 1960), although no relationship was found between actual performance on the task and the increase in restlessness. In the present study the motor restlessness of hyperactive and normal children while seated during the attention task was also examined, using a stabilimetric cushion after the design of Sprague and Toppe (1966).

The third purpose of the study was to investigate the effect of methylphenidate (Ritalin), a central nervous system stimulant, on attention in hyperactive children. A number of investigators have found this drug particularly efficacious in the treatment of hyperactive children (Campbell, Douglas and Morgenstern, 1969; Cohen, Douglas and Morgernstern, 1969; Conners, Eisenberg and Sharpe, 1964; Douglas *et al.*, 1969; Eisenberg, Conners and Sharpe, 1965; Knights and Hinton, 1969; Sprague, Barnes and Werry, 1970; Weiss *et al.*, 1970). However, there is still some question as to its specific effects on sustained attention in hyperactive children (Conners *et al.*, 1964).

METHOD

Subjects

Since the reliable and valid measurement of activity has proved to be a difficult task (Cromwell, Baumeister and Hawkins, 1963), no exact definition of "hyperactive child" in terms of a measured quantity of movement or activity has yet been made. Consequently, the definition of such children in the present study was an operational one based on the selection criteria used. For a child to be included in the study, hyperactivity had to be the major complaint. That is, both the child's parents and teacher had to specify the overactivity of the child as their major complaint and the reason for the referral. Furthermore, such hyperactivity had to have been present as a chronic problem since early childhood and be sustained throughout the day. To ensure as homogeneous a group as possible, children diagnosed by the child psychiatrists as psychotic, epileptic or brain-damaged, or whose major-presenting symptom was behaviour disturbance of an emotional nature, were excluded from the study.

Forty children (34 males and 6 females) who met the above criteria were included

70

in the study. The ages of the *Ss* ranged from 5 to 12 yr (mean age = 8 yr; *s.d.* = 1 yr 9 months) and their intelligence quotients (as measured by the Wechsler Intelligence Scale for Children (Wechsler, 1949)) were not less than 80 (mean *I.Q.* = 111; *s.d.* = 11).

A control group of 19 normal children, matched for age, sex and *I.Q.* with 19 of the hyperactive children, was also tested. The two groups did not differ on mean age ($t = 0{\cdot}003$, $df = 18$, *NS*) or *I.Q.* ($t = 0{\cdot}018$, $df = 18$, *NS*). The control children were selected from the normal school population of the City of Montreal.

All of the children were English speaking, living at home with at least one parent and were attending regular school classes.

Measures

Attention task. The Continuous Performance Test (C.P.T.) is an experimenter-paced task, that is *E* controls the arrival and duration of the task stimuli. A series of letters are presented one at a time on a screen. *S* is required to monitor the screen and respond whenever a specified stimulus, the significant stimulus, appears. Three variables were manipulated in the present study, namely task stimuli, distraction, and interstimulus interval.

There were three separate series of task stimuli, referred to as the *X* the *AX* and the Form sequences. These task stimuli differed in the following respects. In the *X* sequence 12 letters appeared on the screen in a random serial order, the letter *X* being the significant stimulus to which *S* had to respond. In the *AX* sequence the same 12 letters as in the *X* sequence were used, but here the significant stimulus to which *S* responded was the letter *X* only when it was immediately preceded by the letter *A*. The stimuli for the Form sequence consisted of seven geometric shapes in one of five colours. The significant stimulus to which *S* was required to respond was a red triangle.

Each of the three sequences (*X*, *AX* and Form) were presented under two conditions of distraction, minimal and intermittent noise. In the minimal condition *S*'s room was kept as free as possible of extraneous distracting noise. In the intermittent condition white noise (80 decibels) was piped into *S*'s room at frequent random intervals.

Each of the three sequences was displayed at two interstimulus intervals, 1·0 and 1·5 sec. Thus, for both groups of *Ss* there were three sequences (*X*, *AX* and Form), two distraction conditions (minimal and intermittent noise) and two interstimulus intervals (1·0 and 1·5 sec), making 12 trials in all for each *S* (i.e. a four-way complete factorial design).

On each of the 12 trials a total of 200 stimuli were presented. Thus, at the 1·0 sec interstimulus interval a trial was continuous for 3·3 min and at the 1·5 sec interstimulus interval for 5 min.

Motor restlessness. Restlessness while seated was measured with a stabilimetric cushion (Sprague *et al.*, 1966). Movements of the *S* on the seat in the left–right and front–back directions activated microswitches placed underneath the cushion. The microswitches were connected to digital counters which gave a total score based on the movements of the *S* in all directions.

71

Procedure

All *S*s were tested individually. Each *S* visited the hospital on two occasions to complete the initial (pre-drug) testing. During the first session 6 of the 12 trials were given, the remaining 6 trials being given during the second session. The trials were randomized such that order effects were controlled. During the actual testing *S* was alone in the test room, while *E* was in an adjoining room monitoring the equipment and *S*'s behaviour through a one-way screen.

Upon completion of the initial testing each of the 40 hyperactive children was assigned, using a randomized code provided by the Ciba Pharmaceutical Company, to one of two groups, Active Drug or Placebo, using the double-blind technique. The two psychiatrists then titrated the drug for each of their patients until the optimum clinical effect, based on the parents' report of the child's behaviour and the psychiatrist's evaluation of that report, was reached. For most of the children, this optimum clinical effect was reached when the dosage was in the region of 30–40 mg per day. Each hyperactive child returned for retesting (again for two sessions) while on either the active drug or a placebo, 5–7 weeks after completion of the initial testing. The code was not broken until on-drug testing was completed. The control *S*s were seen only for the initial testing (i.e. for only two of the four sessions).

RESULTS

On the C.P.T. there were two dependent measures, the absolute score and the error score. The absolute score is a measure of the *S*s accuracy in detecting the significant stimuli (i.e. No. correct responses/Total no. significant stimuli presented × 100) and the error score is the number of responses to non-significant stimuli.

Hyperactive-control comparisons

Two separate four-way analyses of variance, repeated measures design (Winer, 1962) were used to compare the absolute scores (Table 1a) and the error scores (Table 1b) of the hyperactive and normal children. The four factors were, (1) Groups—hyperactive vs. control, (2) Sequence—X, *AX* and Form, (3) Interstimulus interval—1·0 sec vs. 1·5 sec, and (4) Distraction—noise vs. no noise.

A significant Groups factor ($p < 0.025$) was obtained when the absolute but not the error scores were analysed. This finding indicated that the hyperactive children made significantly fewer correct responses than did the control children (mean hyperactive group = 75·54 per cent, mean control group = 84·49 per cent). However, the groups did not differ with respect to overall number of errors (mean hyperactive group = 6·69, mean control group = 6·12).

Two other main factors were also found to be significant, namely Interstimulus Interval and Sequence. The significant interval factor indicated that both groups of children (hyperactive and control) made more correct responses and fewer errors when the slow (1·5 sec) rather than the fast (1·0 sec) interstimulus interval was used (1·5 sec interval: mean correct both groups = 85 per cent, mean errors = 5·57: 1·0 sec interval: mean correct = 74·5 per cent, mean errors = 7·73).

A main effect for Sequence was found only when absolute scores were analysed and indicated that both groups of children made more correct responses for the X

TABLE 1. TWO FOUR-WAY ANALYSES OF VARIANCE (REPEATED MEASURES) OF THE ABSOLUTE AND ERROR SCORES OBTAINED BY HYPERACTIVE AND NORMAL CHILDREN ON THE CONTINUOUS PERFORMANCE TEST

Source	*df*	(1a) Absolute score *MS*	*F*	*p*	(1b) Error score *MS*	*F*	*p*
Between *S*s							
Groups	1	9126·32	6·76	< 0·025	37·06	0·25	
S (within groups)	36	1350·33			147·80		
Within *S*s							
Sequence	2	903·37	3·64	< 0·05	32·00	1·21	
Groups × sequence	2	363·14			25·04	0·95	
Sequence *S* (within groups)	72	247·95			26·35		
Distraction	1	1·98			8·98	1·30	
Groups × distraction	1	3·51			0·71	0·10	
Distraction *S* (within groups)	36	35·66			6·93		
Interstimulus interval	1	13882·15	158·14	< 0·005	313·34	20·39	< 0·005
Groups × interval	1	0·00			130·56	8·49	< 0·01
Interval *S* (within groups)	36	87·78			15·37		
Sequence × distraction	2	6·64			10·86	1·55	
Groups × sequence × distraction	2	22·40			14·42	2·06	
Sequence × distraction *S* (within groups)	72	33·44			7·01		
Sequence × interval	2	318·03	3·96	< 0·025	15·13	0·92	
Groups × sequence × interval	2	97·65			1·23	0·08	
Sequence × interval *S* (within groups)	72	80·22			16·37		
Distraction × interval	1	94·87			0·43	0·04	
Groups × distraction × interval	1	72·64			0·00	0·00	
Distraction × interval *S* (within groups)	36	49·88			10·25		
Sequence × distraction × interval	2	27·52			1·49	0·19	
Groups × sequence × distraction × interval	2	43·24			0·60	0·08	
Sequence × distraction × interval *S* (within groups)	72	54·83			7·72		

73

(mean = 82·29 per cent) than for the Form (mean = 80·31 per cent) and *AX* (mean = 77·44 per cent) sequences respectively. Mean errors did not differ significantly across the three sequences (mean *X* sequence = 6·74, mean Form sequence = 5·88, mean *AX* sequence = 6·59).

The main effect for Distraction did not reach significance when either absolute or error scores were anlysed. Thus it may be concluded that the distraction factor did not produce any significant decrement in the C.P.T. performance of either group.

Only two interactions reached significance, Groups by Interstimulus Interval (for error scores only) and Sequence by Interstimulus Interval (absolute scores only). The Groups by Interval Interaction is represented in Fig. 1, and indicates that although the hyperactive and control children made a similar number of errors at the fast interval (1·0 sec), the control children made significantly fewer errors at the slower interval (1·5 sec) than the hyperactive children.

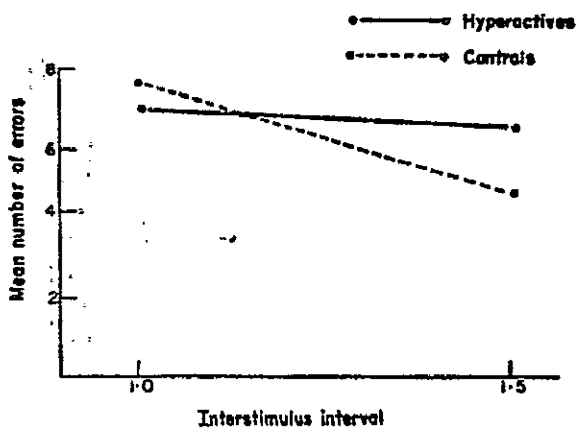


FIG. 1. Incidence of errors on the continuous performance test of the hyperactive and control groups at two interstimulus intervals.

The significant Sequence by Interval interaction indicated that the difficulty of the three sequences, *X*, *AX*, and Form depended on the interstimulus interval used. For the 1·0 sec interval more correct responses were made by both groups to the *X* (mean = 78·36) than to the Form (mean = 74·46) or *AX* (mean = 70·67) sequences. When the 1·5 sec interval was used the order was Form (mean = 86·16), *AX* (84·21) and *X* (82·22).

Motor restlessness

The total activity scores for each *S* accumulated during each of the two testing sessions were first transformed to logarithms to reduce the large variation between *S*s (Winer, 1962) and then analysed by means of a two-way analysis of variance, repeated measures design. The two factors were Groups (hyperactive vs. control) and Session (1st vs. 2nd visit).

The hyperactive children accumulated significantly greater activity scores than the control children (F for groups = 6·84, df = 1/36, $p < 0{\cdot}05$). The main effect for sessions also reached significance, indicating that for both groups restlessness increased during the second session (F = 192·27, df = 1/36, $p < 0{\cdot}001$). Moreover, the significant Groups by Session interaction (F = 7·81, df = 1/36, $p < 0{\cdot}01$) indicated that the restlessness of the hyperactive children increased at a faster rate from the 1st to the 2nd session than did that of the control children (Fig. 2).

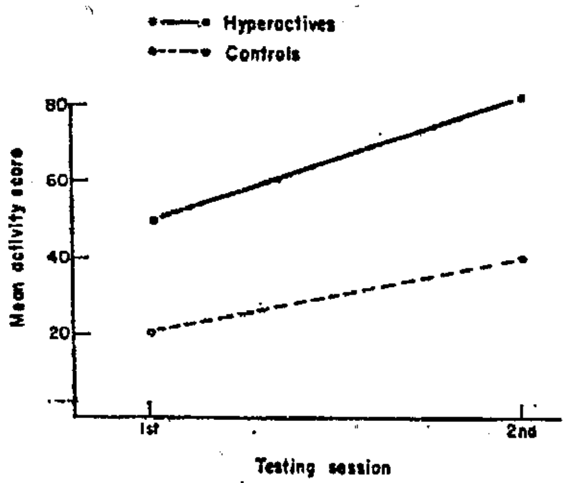


FIG. 2. Restlessness scores of hyperactive and control children during two testing sessions.

Drug comparisons

Despite the fact that the hyperactive children were randomly assigned to either the Active Drug or Placebo group (N = 20 per group) and a double-blind technique was used, it was found that those hyperactive children receiving the active drug had actually obtained poorer pre-drug C.P.T. scores than had those placed on a placebo. Consequently, it was necessary to use analyses of covariance (covarying for initial C.P.T. performance) on both absolute and error scores when evaluating the effect of methylphenidate compared to placebo.

Two-way analyses of covariance (Winer, 1962) were completed on each of the three task sequences (X, AX, and Form) separately. The two factors were Groups (drug vs. placebo) and Interval (1·0 sec vs. 1·5 sec). Distraction (noise vs. no noise) was not used as a main factor in these analyses since the previous analyses (Table 1A and 1B indicated that it did not affect performance.

For all three task sequences, the active drug group had a significantly higher absolute score than did the placebo group ($p < 0{\cdot}005$ in all three cases). The active drug group also made significantly fewer errors than the placebo group on the X ($p < 0{\cdot}005$) and Form ($p < 0{\cdot}05$) but not the AX sequences.

The main effect for Interval was significant for all three task sequences when absolute scores were analysed. As in the initial analyses, this finding indicated that

75

*S*s made more correct responses when the slow rather than the fast interstimulus interval was used. They also made fewer errors when the slow interval was used, this effect reaching significance only in the analysis of error scores on the *AX* task.

The Groups by Interval interaction was not significant in the analyses of absolute scores for any of the three tasks. When error scores were analysed, however, this interaction reached significance in the *X* condition ($p < 0{\cdot}5$) and showed a trend in the *AX* condition ($p < 0{\cdot}10$). This interaction is illustrated in Fig. 3 and indicates that whereas the active drug group made fewer errors at the slow than at the fast interstimulus interval, the placebo group made a similar number of errors at both speeds. The behaviour of the active drug group in this respect is reminiscent of the behaviour of the normal control children at the slower speed (see Fig. 1).

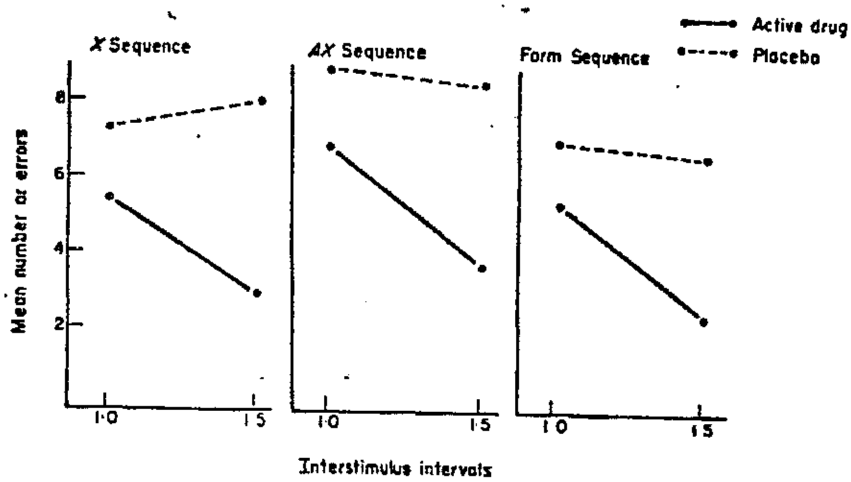


FIG. 3. Incidence of errors on the continuous performance test of the active drug and placebo groups at two interstimulus intervals.

DISCUSSION

The results of the present study lend support to the hypotheses that hyperactive children are deficient with respect to their ability to maintain attention to a task and are physically more restless than normal children.

On a three-part task (*X*, *AX* and Form) requiring the monitoring of a screen, the hyperactive children detected significantly fewer of the significant stimuli. They also made significantly more responses to non-significant stimuli than did the control children at the slow, but not at the fast interstimulus interval. It should be noted, however, that the performance of the hyperactive children was not as impaired as that of the brain-damaged and centrencephalic epileptic children studied by Rosvold *et al.* (1956) and Fedio and Mirsky (1969).

Three task variables were manipulated in an attempt to ascertain whether hyperactive and normal children would respond differently under different conditions. The three variables were (1) sequence, (2) distraction and (3) interstimulus interval.

It was found that the three sequences used did not differ in order of difficulty for the hyperactive and normal children, all subjects finding the *X* sequence easier than the Form or *AX* sequences at the slow (1·5 sec) interstimulus interval and the Form sequence easier than the *AX* or *X* sequences at the fast (1·0sec) interval (as measured by the absolute score).

Similarly, the second variable, namely distraction, did not differentially affect hyperactive and control children, contrary to the widely held view that hyperactive children are more susceptible to distraction than are normal children. The present study found that the performance of neither group of children was affected by the particular type of distraction used (intermittent white noise). Fedio *et al.* (1969) also failed to find any effect of an intermittent pure tone of 101 decibels on the C.P.T. performance of normal and epileptic children. Sen and Clarke (1968) point out that susceptibility to distraction is related to task difficulty. The more difficult the task, especially as it reaches the limits of the information processing capacity of the individual, the more deleterious the effect on performance of extraneous stimuli in the form of a distractor. It is possible that if the demands placed on the information processing capacities of the individual by the C.P.T. had been greater, then the intermittent noise might have produced a decline in performance. Certainly, the hypothesis that hyperactive children are more susceptible to distraction than normal children is neither confirmed nor denied by the findings of the present study but needs to be tested more systematically, taking into account such variables as type of distractor and level of task difficulty (i.e. information load).

The third variable manipulated, namely interstimulus interval, did affect the C.P.T. performance (error scores only) of hyperactive and normal children differentially, since normal children made significantly fewer errors than the hyperactive children when a slow (1·5 sec) rather than a fast (1·0 sec) speed of presentation was used. Since "errors" are responses to non-significant stimuli, this finding suggests that additional time between stimuli helped normal children more than hyperactive children to evaluate each stimulus more efficiently and thus to inhibit their responses to non-significant stimuli. That hyperactive children are more prone than normal children to impulsive responding is well documented (Campbell *et al.*, 1969; Cohen *et al.*, 1969; Conners and Greenfeld, 1966; Stevens *et al.*, 1968) and it would be interesting in a future study to see if an even longer interstimulus interval would help hyperactive children to make fewer errors on the C.P.T.

As expected, the hyperactive children were more restless (as measured by a stabilimetric cushion activity score) than the normal children. Also, their restlessness increased from the 1st to the 2nd testing session at a faster rate than did that of the normal children. While Cromwell *et al.* (1963) have suggested that the "hyperactivity" or restlessness of these children is a result of an inability to maintain attention, it is not yet possible to specify the cause-effect relationship, if any, between these two variables. As techniques for measuring activity or movement become more sophisticated, it may be possible to examine in detail the relationship between these behaviours.

The findings of the present study demonstrate that methylphenidate significantly improves the C.P.T. performance of hyperactive children: the hyperactive children on the active drug detected more significant stimuli and made fewer errors or (impulsive) responses to non-significant stimuli than did the hyperactive children given a placebo. In fact, the Drug Groups by Interstimulus Interval interactions obtained when *error* scores for the X ($p < 0{\cdot}05$) and AX ($p < 0{\cdot}10$) sequences were analysed suggests that the effect of the drug is to enable those taking it to make fewer errors at the slow speed compared to the fast speed of presentation of stimuli. In other words, it apparently helped the hyperactive children to evaluate the stimuli more effectively and thus to inhibit their responses to non-significant stimuli. This is in contrast to the hyperactive children given a placebo who did not show a similar reduction in errors at the slow interstimulus interval.

An important question which the present study could not answer owing to technical difficulties was whether methylphenidate reduced the physical restlessness of the hyperactive children. Such a reduction in the physical restlessness of hyperactive, emotionally disturbed boys given methylphenidate has recently been reported by Sprague *et al.* (1970). Clinical observation by the present investigators and the parents of the hyperactive children suggests that methylphenidate did reduce restlessness (as found by Knights *et al.*, 1969; Weiss *et al.*, 1970). In the light of the Cromwell *et al.* (1963) suggestion that the restlessness of hyperactive children is not mere activity *per se*, but continued shifting from task to task because of short attention span, it may be that the effect of methylphenidate is to improve the ability of hyperactive children to pay attention (as demonstrated on the C.P.T. in the present study) such that they shift less from task to task and thus appear less restless or active.

In conclusion, while the present study provides some evidence that hyperactive children are deficient with respect to sustained attention it is clear that future research must examine behaviours that are more closely related to the actual attentional demands placed on the child in the school setting. In other words, the emphasis must be on the discovery of the variables influencing attention in the school setting and on the manipulation of these variables in an attempt to ameliorate the attentional difficulties currently experienced in school by many hyperactive children. Also, while the present study provides some evidence that methylphenidate has a beneficial effect on the performance of certain relatively simple tasks over a short period of time, we need to know whether this beneficial effect also applies to more complex behaviours over longer periods of time.

SUMMARY

The maintenance of attention to an experimenter-paced task requiring the detection of significant stimuli was impaired in hyperactive children. When compared with a matched normal control group, the hyperactive children detected fewer of the significant stimuli and made more incorrect responses to non-significant stimuli. The presence or absence of an auditory distractor had no influence on the performance of either group of children. Those hyperactive children treated with methylphenidate (ritalin) showed a significant improvement in all aspects of their performance when compared to a control group of hyperactive children given a placebo.

REFERENCES

BAKER, C. H. (1960) Observing behavior in a vigilance task. *Science* **132**, 674–675.

CAMBELL, S., DOUGLAS, V. and MORGENSTERN, G. (1969) Cognitive styles in hyperactive children and the effect of methylphenidate. Paper read at Canadian Psychological Association Annual Meeting, Toronto.

COHEN, N., DOUGLAS, V. and MORGENSTERN, G. (1969) Psychophysiological concomitants of hyperactivity in children. Paper read at Canadian Psychological Association Annual Meeting, Toronto.

CONNERS, C. K., EISENBERG, L. and SHARPE, L. (1964) Effects of methylphenidate (ritalin) on paired-associate learning and Porteus Maze Performance in emotionally disturbed children. *J. Consult. Psychol.* **28**, 14–22.

CONNERS, C. K. and GREENFELD, D. (1966) Habituation of motor startle in anxious and restless children. *J. Child Psychol. Psychiat.* **7**, 125–132.

CROMWELL, R. L., BAUMEISTER, A. and HAWKINS, W. F. (1963) Research in activity level. In *Handbook of Mental Deficiency* (Edited by ELLIS, N. R.), McGraw-Hill, New York.

DOUGLAS, V., WEISS, G. and MINDE, K. (1969) Learning disabilities in hyperactive children and the effect of methylphenidate. Paper read at Canadian Psychological Association Annual Meeting, Toronto.

EISENBERG, L., CONNERS, C. K. and SHARPE, L. (1965) A controlled study of the differential application of outpatient psychiatric treatment for children. *Japan. J. Child Psychiat.* **6**, 125–132.

FEDIO, P. and MIRSKY, A. F. (1969) Selective intellectual deficits in children with temporal lobe or centrencephalic epilepsy. *Neuropsychologia* **7**, 287–300.

KNIGHTS, R. M. and HINTON, G. (1969) The effects of methylphenidate (ritalin) on the motor skills and behaviour of children with learning problems. *J. Nerv. Ment. Dis.* **148**, 643–653.

LAPOUSE, R. and MONK, M. A. (1958) An epidemiologic study of behavior characteristics in children. *Am. J. Publ. Hlth* **48**, 1134–1144.

ROSVOLD, H. E., MIRSKY, A. F., SARASON, I., BRANSOME, E. D. and BECK, L. H. (1956) A continuous performance test of brain damage. *J. Consult. Psychol.* **20**, 343–350.

SCHULMAN, J. L., KASPAR, J. C. and THRONE, F. M. (1965) Brain damage and behavior. *A Clinical-Experimental Study*. C. Thomas, Springfield, Illinois.

SEN, A. and CLARKE, A. M. (1968) Some factors affecting distractibility in the mental retardate. *Am. J. Ment. Defic.* **73**, 50–60.

SPRAGUE, R. L., BARNES, K. R. and WERRY, J. S. (1970) Methylphenidate and thioridazine: learning, reaction time, activity, and classroom behavior in disturbed children. *Am. J. Orthopsychiat.* **40**, 615–628.

SPRAGUE, R. L. and TOPPE, L. K. (1966) Relationship between activity level and delay of reinforcement in the retarded. *J. Experiment. Child Psychol.* **3**, 390–397.

STEVENS, D. A., BOYDSTUN, J. A., ACKERMAN, P. T. and DYKMAN, R. A. (1968) Reaction time, impulsivity, and automonic lability in children with minimal brain dysfunction. *Proc. 76th Ann. Conv., Am. Psychol. Assoc.*

STEWART, M. A., PITTS, F. N., CRAIG, A. G. and DIERUF, W. (1966) The hyperactive child syndrome. *Am. J. Orthopsychiat.* **36**, 861–867.

WECHSLER, D. (1949) *Wechsler Intelligence Scale for Children*. Psychological Corporation, New York.

WEISS, G., MINDE, K., DOUGLAS, V., WERRY, J. S. and SYKES, D. H. (1970) A comparison of the effects of chlorpromazine, dextroamphetamine and methylphenidate on the behaviour and intellectual functioning of the hyperactive child. *Can. Med. Assoc. J.* In press.

WERRY, J. S. and SPRAGUE, R. L. (1970) Hyperactivity. In *Symptoms of Psychopathology* (Edited by COSTELLO, C. G.), Wiley, New York.

WINER, B. J. (1962) *Statistical Principles in Experimental Design*. McGraw-Hill, New York.

79

Journal of Abnormal Child Psychology, Vol. 11, No. 4, 1983, pp. 487-496

1250

The Computerized Continuous Performance Task: A New Measure of Inattention

Steven H. Klee
Beth Israel Medical Center

Barry D. Garfinkel
University of Minnesota Medical School

Attentional performance was measured using a computerized continuous performance task, several psychometric tasks, and ratings of classroom behavior. Subjects were 51 children in the inpatient and day hospital programs of a psychiatric hospital. The relationship between performance on the computerized task and all other measures was examined. Results indicated that the continuous performance task significantly correlated with several other psychometric measures of inattention, ratings of inattention, impulsivity, and hyperactivity. The CPT had slightly better sensitivity and the same specificity as the Conners Teacher Rating Scale in identifying Conduct and Attention Deficit Disordered children. Implications for the use of the computerized continuous performance task as a screening measure for attentional difficulties is discussed.

The Continuous Performance Task (Rosvold, Mirsky, Sarason, Bransome, & Beck, 1956) was originally designed to detect and study brain damage in both children and adults. This early report indicated that the brain-damaged group's performance was consistently inferior when compared to nonimpaired controls. It was suggested that the inferior performance resulted from decreased alertness in the brain-damaged group.

Vigilance or continuous performance tasks have been used to measure sustained attention in adults (Kupietz & Richardson, 1978). Their use with Attention Deficit Disorder (ADD) children is also appropriate as these children are described as inattentive, distractible, fidgety, restless, and impulsive (Cantwell, 1972). When vigilance performance of normal

Manuscript received in final form February 23, 1983.

children was compared to that of ADD children, the index children performed less accurately (Sykes, Douglas, & Morgenstern, 1973). These authors concluded that the clinical group's poor performance was a symptom of generalized behvioral problems that appeared as inattentiveness and distractibility in other settings. Kupietz and Richardson (1978) reported that teacher ratings of classroom behavior were significantly and negatively correlated with visual vigilance task performance. Additional evidence for the relationship of poor vigilance performance and attention deficits has been provided by Douglas (1972) and more recently by Weingartner, Rapoport, Buchsbaum, Bunney, Ebert, Mikkelsen, and Caine (1980).

As continuous performance tasks are used more frequently in the investigation of attention deficits, it will become increasingly important to clarify the behavioral, perceptual, and cognitive dimensions measured by the Continuous Performance Test (CPT). Recent studies employing the CPT have replaced the laboratory equipment with more precise microcomputers. The computer is able to present stimuli at clearly defined intervals and simultaneously record subject responses. The CPT provides two dependent measures: omission and commission errors. Omission errors are reportedly sensitive to problems of inattention, while commission errors are thought to be indicative of problems of impulsivity (Sostek, Buchsbaum, & Rapoport, 1980). However, few studies have attempted to correlate CPT performance and other commonly used measures of inattention, concentration, reflectivity, impulsivity, and childhood behavior.

The present study explored the relationship between a CPT task and other psychometric measures for children hospitalized with psychiatric disorders. Vigilance as measured by CPT performance was hypothesized to correlate strongly with other psychometric measures of attention, reflectivity, and impulsivity. CPT performance was also compared to ratings of the child's behavior on the Conners Teacher Rating Scale (CTRS; Conners, 1969). It was hypothesized that poor CPT performance would identify children with marked behavioral problems as indicated by high overall scores on the CTRS. More specifically, it was predicted that CPT performance would correlated significantly with CTRS Factor 1 (conduct problems), Factor 2 (inattention), and Factor 4 (hyperactivity), and not with Factor 3 (anxiety) and Factor 5 (sociability). If the computerized CPT could identify a highly distractible and impulsive group of children, this would be a brief and accurate method of validating teacher observations and would serve to confirm deficits measured by other psychometric tests. A secondary purpose of this study was to explore the relationship of the CPT to specific diagnostic categories. Of particular interest was the usefulness of

the CPT alone and in combination with other measures in providing further clinical description of Conduct and Attention Deficit Disorder children.

METHOD

Subjects

All children admitted to the full or partial hospitalization programs at a children's psychiatric hospital constituted the subject pool. Exclusionary criteria included documented seizure disorders, brain damage, and IQ scores on the WISC-R (Wechsler, 1974) of less than 70. IQ scores were coded on a 5-point scale ranging from 1 (borderline IQ = 70-79) to 5 (superior IQ = 120+). These categories matched those provided in the WISC-R manual. The mean IQ of the subject sample was 2.7 (SD = .9) (low average-average range). The final sample consisted of 51 subjects (13 F, 38 M) with a mean age of 12.5 years (range: 7.9-16.5).

Patients involved in this study were referred for a wide variety of presenting problems, including delinquent behavior, mood disturbances, severe behavioral problems (at school and at home), and thought disorders. On the basis of DSM III diagnostic categories, subjects' admission diagnoses fell into five major categories, including Conduct Disorders (N = 23), Schizophrenic Disorders (N = 10), Attention Deficit Disorders (N = 7), Major Affective Disorders (N = 5), and Other (N = 6).

Procedure

All subjects were tested individually by one of two experimenters for 30 minutes and included the following measures:

Continuous Performance Task. This task was similar in concept to that employed by Rosvold et al. (1956). The CPT was programmed in PASCAL utilizing an Apple II computer. The task consisted of 10 letters flashing on the center of a black-and-white video monitor for 130 milliseconds at the rate of 600 milliseconds between letters. The target was the letter *S* followed by the letter *T*. The task was to press a bar when the target appeared and to avoid responding to other letter combinations. All subjects were exposed to 500 letters, with 10% of them being targets. Two dependent measures were obtained. An error of omission was scored for each target missed, while a commission error was scored when a response was made to a nontarget stimulus.

Kagan Matching Familiar Figures Test (MFFT). The children's form of the MFFT was used. This test measures reflectivity-impulsivity (Kagan, 1964), discriminates between ADD and non-ADD children, and is sensitive to stimulant drug effects (Campbell, Douglas, & Morgenstern, 1971). The test consists of 12 pictures of familiar objects followed by six variants of each picture. The subject selects from six choices the one identical to the original picture. Scoring includes latency of first response and total number of errors.

WISC-R Subtests. The Coding and Arithmetic subtests from the Wechsler Intelligence Scale for Children-Revised were administered (Wechsler, 1974). Performance on both subscales can be influenced by attention span and distractibility (Kaufman, 1979).

Behavioral Ratings. Child care workers completed the Conners Teacher Rating Scales (CTRS) on all of the children in their units. Each child care worker assigned to a particular child had the opportunity to observe the subject in both a structured classroom setting and a less structured unit setting. In order to compensate for daily fluctuations in workers' observations and subjects' behavior, rating scales were filled out at the end of the day on 3 consecutive days during the middle portion of the week. These scores were then averaged to form one score. The 3 days of rating overlapped with those of psychometric testing. The CTRS has 39 items that were scored from 0 ("not at all") to 3 ("very often"). Thus, higher total scores were indicative of more severe behavioral problems.

RESULTS

Table I shows the means, standard deviations, and ranges on the rating scale and psychometric measures for all subjects. Inspection of Table I shows that there was a wide range of performance on both the psychometric tasks and rating scales. Thus, the measures employed were sensitive in capturing the performance variability among subjects.

Table II details the correlations between CPT performance and all other measures. Correlations for errors of omission and commission were obtained separately as well as in combination (CPTOT). Consistent with Kupietz and Richardson (1978), CPT errors were found to be highly correlated with age. Therefore, partial correlations controlling for age were computed. Omission errors were positively correlated with MFFT errors and child care workers' ratings on the CTRS. High negative correlations were also found between omission errors and latency of first response on the MFFT, and scores on the WISC-R coding and arithmetic subtests. Commission errors were correlated negatively with latency and total response time

Table I. Rating Scale and Psychometric Profile

Variable	N	Mean	SD	Range
CTRS[a]	51	38.3	17.7	9.7–87.0
Omission errors	51	20.1	8.8	3–38
Commission errors	51	20.8	17.0	3–68
MFFT[b] (first response)	51	16.5	10.2	4.6–46.8
MFFT (total time)	51	22.6	11.9	7.8–58.5
MFFT (errors)	51	8.3	5.2	0–23
Coding	51	7.5	3.7	2–16
Arithmetic	51	8.3	2.3	4–13

[a]CTRS = Conners Teacher Rating Scale.
[b]MFFT = Kagan Matching Familiar Figures Test.

on the MFFT, positively with MFFT total errors and with child care workers' CTRS ratings. CPTOT was positively correlated with CTRS scores and MFFT errors, and negatively correlated with MFFT first response time and the coding subtest.

Partial correlations between CPT performance and CTRS factor scores are presented in Table III. All three measures from the computerized task were positively correlated with both Factor 2 (inattention) and Factor 4 (hyperactivity).

Though the above correlations were obtained, the relationship between CPT performance and behavioral ratings required further clarification. Subjects with a high CPT error rate were identified using a cutoff of 1/2 SD above the mean. As CPT and age were significantly correlated, groups were formed for younger (13.0 years and below) and older (13.1 years and above) children. Children with high error rates were chosen through comparison with their appropriate age group means. Fifteen (N = 15) children were selected as having a high error rate.

The percentage of subjects who were rated on each of the 39 behaviors as appearing "often" (2) or "very often" (3) were also noted. No item had

Table II. Correlations Between CPT Task and Other Psychometric Variables

	Partial order controlling for age						
	CTRS[c]	MFFT[d] (1st)	MFFT (total)	MFFT (errors)	Coding	Arithmetic	Age
Omission errors	.32[a]	.28[a]	−.09	.31[a]	−.32[a]	.37[a]	.52[b]
Commission errors	.36[b]	−.35[b]	.33[a]	.34[a]	.25	.12	.49[b]
CPT total errors	.38[b]	−.36[b]	−.26	.28[a]	.31[a]	.15	.26

[a]$p < .05$.
[b]$p < .01$.
[c]CTRS = Conners Teacher Rating Scale.
[d]MFFT = Kagan Matching Familiar Figures Test.

Table III. Correlations Between CPT Task and Conners Teaching Rating Scale Factor Scores

	Factors				
	1 (Conduct)	2 (Inattention)	3 (Anxiety)	4 (Hyperactivity)	5 (Sociability)
Omission errors	−.05	.31[a]	−.10	.36[a]	.21
Commission errors	−.03	.33[a]	−.01	.34[a]	.05
CPT total errors	−.04	.33[a]	−.04	.35[b]	.12

[a] $p < .05$.
[b] $p < .01$.

100% endorsement. The three most endorsed CTRS items for the high CPT error group were "poor frustration tolerance" (67%), "anxious to please" (67%), and "excessive demand for attention" (60%). These items were followed by "inattentive/distractible" (47%), "short attention span" (40%), "sensitive to criticism" (40%), and "acts smart" (40%). Items occurring rarely, and therefore not characteristic of this group, were "does not get along with opposite sex" (0%), "attendance problems" (0%), "cries easily" (7%), and "shy" (7%).

One purpose of the present study was to determine the usefulness of the CPT in aiding with the clinical description and diagnosis of children. Of importance was the relative contribution of the CPT when compared to the CTRS alone. To answer this question, we selected children with scores 1/2 *SD* above the mean for both CPT performance and CTRS ratings. Sensitivity and specificity for each of these measures were determined.

Ideally, it would have been preferable to examine the sensitivity and specificity of these measures in relation to the diagnosis of ADD alone. However, as reported earlier, few subjects were hospitalized with this diagnosis. As a compromise, the ADD and Conduct Disorder groups were pooled. Thirty children met diagnostic criteria for Conduct Disorder and Attention Deficit Disorder, a group of children most likely to show deficits in concentration, impulsivity, and behavioral symptoms. Fourteen had extreme CPT scores yielding a sensitivity, or number of true positives, with this measure of 47%. By comparison, of the 21 children having diagnoses other than Conduct Disorder or Attention Deficit Disorder, 16 were not extreme responders on the CPT. Thus, specificity or number of true negatives for the CPT was 76%. The CPT was more sensitive than the CTRS, which identified 13 of the 30 Conduct Disorder and ADD children, providing a sensitivity of 43%. The specificity of the CTRS was equal to the CPT, with 76% of the other children being properly identified as not ADD or Conduct Disorder.

DISCUSSION

The usefulness of a computerized continuous performance task as a descriptive measure for attentional problems received substantial support. CPT performance was significantly correlated with other psychometric and behavioral ratings of inattention, impulsivity, and hyperactivity. A combination of omission and commission errors was found to be associated with (a) measures of impulsivity, distractibility, and lack of reflection on the MFFT; (b) overall behavioral problems as measured on the CTRS by child care workers; and (c) sustained attention as measured by the coding subtest of the WISC-R. Omission errors alone were found to correlate significantly with the arithmetic subtest, which has been viewed as a measure of sustained attention (Kaufman, 1979). Additionally, commission errors were significantly correlated with MFFT total time, which was felt to reflect the subjects' sustained attention on a task. These findings suggest that the CPT measures cognitive functions similar to other psychometric tests and behavioral ratings.

In general, the CPT was found to correlate strongly with these other measures in general, with omission scores not correlating specifically with sustained attention tasks or commission scores with measures of impulsivity alone. Rather, both of these scores correlated in a more global way with the other measures. Thus, the generally accepted notion that omission errors measure inattention and commission errors measure impulsivity was not supported in this study. Given this, the composite of CPTOT appears to be a useful summary measure in describing a subject's CPT performance.

It was expected that CPT would correlate with Factors 1 (conduct), 2 (inattention), and 4 (hyperactivity) on the CTRS. Only Factors 2 (inattention) and 4 (hyperactivity), however, correlated with CPTOT. These findings support the clinical observations that attention deficits are most often associated with symptoms of hyperactivity, impulsivity, and restlessness rather than generalized behavioral problems.

Identifying specific CTRS items characteristic of the high CPT responders provided additional support for this position. The behavioral items most endorsed by child care staff for extreme CPT responders included many of the common symptoms of Attention Deficit Disorder. Among these were poor frustration tolerance, excessive attentional demands, short attention, and high distractibility. Also, several highly endorsed items could be related to conduct disorders, including sensitivity to criticism, acting smart, and the excessive demand for attention. While CPT performance appears to be related to behaviors associated with Attention Deficit Disorders, this study was not able to demonstrate that poor

CPT performance was unique to this diagnosis. Specifically, the CPT appeared to be correlated with behaviors that might also be associated with the diagnosis of Conduct Disorder.

Limitations of the present study result in part from the sample used—namely, hospitalized children. These children are likely among the most symptomatic children in the community. As such, there is likely a great overlap of symptoms and deficits across a variety of diagnoses. Furthermore, the method of establishing diagnoses as well as the diagnostic heterogeneity of subjects were major limitations. Diagnoses were taken from admission evaluations with no attempt to validate them through independent standardized interviews. Additionally, there was no attempt to gather a minimum number of subjects in certain diagnostic categories. Rather, the current hospital census constituted the total subject pool. Therefore, there was marked overrepresentation of Conduct Disordered diagnoses, while a diagnosis of Attention Deficit Disorder was rare. Such a breakdown prevented any meaningful comparisons between these two diagnostic groups. For the present study, these diagnostic groups were pooled to provide comparisons with all other diagnostic categories.

Alone, both the CTRS and CPT had comparable moderate sensitivity and good specificity. These results indicate that with this hospitalized population, extreme responding either on the CTRS or CPT was related to a diagnosis of Conduct Disorder or ADD. A person identified as an extreme responder on either of these measures was most likely to receive a diagnosis of Conduct Disorder or ADD. Although many receiving such diagnoses were not extreme responders, a person who was either a poor CPT responder or rated poorly on the CTRS would only rarely be diagnosed as anything other than ADD or Conduct Disorder (95% specificity when combined). It should be noted that the CPT provided slightly greater sensitivity than the CTRS. It appears, therefore, to have utility as a classroom screening device with similar application as the CTRS.

The correlational data presented here are necessarily preliminary and speculative due to the nature of this study. In addition, the divisions into extreme groups using *1/2 SD* cutoff were both retrospective and arbitrary. The value of using the CPT as a screening measure for attention deficits needs to be further addressed in future studies. Such studies should specifically address how well the CPT differentiates patients along diagnostic classifications. Comparisons among large groups of ADD, Conduct Disorder, and normal controls are needed.

In summary, the CPT was found to be a brief engaging task that can be applied to a clinical population in the preadolescent age range. The CPT appears to measure inattention and impulsivity in a global way comparable to, but not exactly the same as, the MFFT and the WISC-R Coding and

Arithmetic subtests. Similarly, it appears related to the child care workers' ratings of the child's behavior. When the results of the CPT and CTRS were combined, they yielded a good deal of diagnostic certainty. Futher study of specific behaviors that appear highly correlated to CPT performance would greatly assist teacher and clinician understanding of the child experiencing attention difficulties in the classroom. Classroom behaviors such as daydreaming, fidgeting, out-of-seat behavior, and restlessness may be better understood in the context of existing cognitive deficits as measured by this task. This study lent support to the observation that concentration and attentional problems can be easily detected on a brief screening device such as a CPT, and that these deficits are generalized and recognized by other psychometric tasks as well as behavioral ratings.

REFERENCES

Campbell, S. B., Douglas, V. I., & Morgenstern, G. Cognitive styles in hyperactive children and the effects of methylpheuidate. *Journal of Child Psychology and Psychiatry*, 1971, *12*, 56-67.

Cantwell, D. P. Psychiatric illness in the families of hyperactive children. *Journal of Child Psychology and Psychiatry*, 1972, *19*, 145-153.

Conners, C. K. A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry*, 1969, *126*, 884-888.

Douglas, V. I. Stop, look and listen. The problem of sustained attention and impulse control in hyperactive and normal children. *Canadian Journal of Behavioral Science*, 1972, *4*, 259-282.

Kagan, J. *The matching familiar figures test*. Cambridge: Harvard University, 1964.

Kaufman, A. S. *Intelligent testing with the WISC-R*. New York: Wiley, 1979.

Kupietz, S. S., & Richardson, E. Children's vigilance performance and inattentiveness in the classroom. *Journal of Child Psychology and Psychiatry*, 1978, *19*, 145-153.

Rosvold, H., Mirsky, A., Sarason, I., Bransome, E., & Beck, L. A continuous performance test of brain damage. *Journal of Consulting Psychology*, 1956, *20*, 343-350.

Sostek, A. J., Buchsbaum, M. S., & Rapoport, J. L. Effects of amphetamine on vigilance performance in normal and hyperactive children. *Journal of Abnormal Child Psychology*, 1980, *8*, 491-500.

Sykes, D. H., Douglas, V. I., & Morgenstern, G. L. Sustained attention in hyperactive children. *Journal of Child Psychology and Psychiatry*, 1973, *14*, 213-220.

Wechsler, D. *Wechsler intelligence scale for children-revised*. New York: Psychological Corporation, 1974.

Weingartner, H., Rapoport, J. L., Buchsbaum, M. S., Bunney, W. E., Ebert, M. H., Mikkelsen, E. J., & Caine, E. D. Cognitive processes in normal and hyperactive children and their response to amphetamine treatment. *Journal of Abnormal Child Psychology*, 1980, *5*, 187-197.

The MCA

The MCA is a clinical research tool that has been used with considerable success to assess the effects of methylphenidate (Ritalin) in the treatment of attention deficit disorders. Although the MCA is not a proven diagnostic instrument, it has been useful when combined with classroom behavior ratings, such as the CPTQ-A and HRS, and other diagnostic procedures to help to determine whether a child might have an attention deficit disorder.

The MCA is a 25 minute psychophysiologic test using visual stimuli to measure attentional variables. This instrument has documented test-retest reliability and has been standardized for 6 to 12 year olds. The MCA is an adaptation of the VIRTEST (see Yellin, A.: A standard visual stimulus for use in studies of attention and attention deficit disorders. Res. Commun. Psychol. and Psychiat. Behav., 1980; 5:137-143 and Yellin, A., Hopwood, J. and Greenberg, L.: Adults and adolescents with attention deficit disorder. J. Clin. Psychopharm., 1982; 2:133-136).

Programmed for use with the Apple IIe, the MCA presents two easily discriminated visual stimuli in two test conditions: one in which the signal occurs infrequently in comparison to the nonsignal; in the second condition, the signal occurs more frequently than the nonsignal. The former condition, or signal infrequent, is similar to the traditional vigilance test in which errors of omission (or not responding to the signal) can be interpretd as a measure of inattentiveness. In the signal frequent condition which is a unique feature of the MCA, errors of commission (incorrectly responding to the nonsignal) can be interpreted as a measure of impulsivity or difficulty inhibiting inappropriate responses. In addition to the two types of errors, the MCA also measures response times and standard deviations which can be interpreted as an index of variability. Errors of omission, response time and standard deviations have been reported to improve when psychostimulants are administered to individuals with attention deficit disorder.

Lawrence M. Greenberg
Department of Psychiatry
University of Minnesota
Box 95
420 Delaware Street, SE
Minneapolis, MN 55455
(612) 373-8871

APPLIED INNOVATIONS, INC.

IN-SITE PSYCHOLOGICAL TESTING SOFTWARE

... for Health Care and Human Service Professionals

86

OTHER PSYCHOLOGICAL TESTING SOFTWARE AVAILABLE FROM ai

NEUROPSYCHOLOGICAL TESTING

SBORDONE-HALL MEMORY BATTERY (Revised & Updated) – provides fully automatic serial testing of 18 discrete memory functions utilizing computer generated test stimuli.

- Produces a comprehensive quantitative and qualitative analysis of patient's performance in comparison to normative data.
- Personalized testing, understandable by all levels of users with little or no supervision required.
- Excellent for clincial assessment, cognitive rehabilitation or research.

Available on Apple II+, IIe, IIc, III (Item Number S1000) $375

THE DIGIT-DIGIT ATTENTION TEST – fully automatic serial testing of complex attention skills utilizing randomly generated test stimuli.

- Program actually trains the individual *how* to take the test.
- Program maintains up to 900 subject files (bibliographic data, test scores, histograms of patient performance, etc.).
- Program generates various statistical and trend analyses.

Available on Apple II+, IIe, IIc, III (Item Number S1010) $250

COGNITIVE REHABILITATION

COMPLEX-ATTENTION REHABILITATION – provides training in attentional tasks utilizing visual tracking of single or multiple stimuli.

- Interacts directly with the patient, determining the task and level of complexity.
- Advanced levels to utilize language-mediated complex attentional skills.

Available on Apple II+, II e, IIc (Item Number S1020) $195

PROBLEM-SOLVING REHABILITATION I – provides training in visual-spatial problem-solving skills through use of a wide range of systematic tasks.
Available on Apple II+, IIe, IIc (Item Number S1030) $195.00

PROBLEM-SOLVING REHABILITATION II – improves planning skills to facilitate effective problem-solving behavior and trains the patient to plan and evaluate the consequences of future behavior.
Available on Apple IIe, IIc (Item Number S1040) $195.00

These programs will be available soon for IBM PC and compatibles. For more information or to order, write or call us toll-free at 1-800-272-2250.

ai

87

Applied Innovations, Inc., South Kingstown Office Park
Wakefield, Rhode Island 02879, (401) 789-5081, (800) 272-2250

88

THE CONTINUOUS PERFORMANCE TASK

(Version 3.1)

***** AN OPERATING MANUAL *****

Division of Child and Adolescent
Psychiatry
University of Minnesota Medical School
University Hospitals
Box 95 Mayo Memorial Building
420 Delaware Street S.E.
Minneapolis, Minnesota 55455
(612) 373-8871

89

TABLE OF CONTENTS

90

INTRODUCTION

This manual is a description of the computerized Continuous Performance Task (CPT). It is designed to assist in the administration of the CPT and provides a basic description of its program. Every attempt has been made to create a "user-friendly" system, i.e., that which requires minimal computer (specifically Apple IIe) knowledge. One will find, however, that such knowledge will aid in the understanding of the program's structure and logic. Some commonly used computer terms will be used to facilitate the brevity and clarity of this manual. It would, therefore, be advantageous to familiarize oneself wth basic computer and Apple IIe terminology.

Many users too often attempt to operate a program without first reading the accompanying manual. We recommend this not be attempted with the CPT program, since many errors will result and could lead to inappropriate parameters being set and the generation of invalid and unreliable results.

HARDWARE

The CPT software is written for specific hardware. Little deviation from these requirements will be tolerated by the software. For consistency sake, it is best to adhere to the following:

1. Apple IIe with two (2) disk drives -- non-enhanced version of ROM and microprocessor (slot #6);
2. Printer with parallel interface card (slot #1);
3. Composite color video monitor (diagonal diameter should be approximately 11");
4. 80-column card (non-extended memory; auxiliary slot #3) This card is optional but recommended (see footnote);
5. Mountain Hardware Inc. THE CLOCK™ (slot #5)
   a) The interrupt frequency required is 10 milliseconds. Necessary modifications can be performed by an individual familiar with electronics and soldering with a 25 watt soldering gun. See THE CLOCK™ operating manual (#11-00229-04) page 34;
   b) The switch on THE CLOCK™ must be in the WRITE position (see operating manual, page 3);
6. Apple joystick;
7. DOS 3.3 initialized disks (for data storage).

Footnote: Although not essential for operation of the CPT, the 80-column card allows for results to be displayed on the video-monitor in a readable fashion.

TESTING ENVIRONMENT

To obtain valid and reliable results, special care should be taken to insure a suitable testing environment. The computer should be arranged so that the center of the video monitor is at or slightly below eye level. The keyboard should be arranged below the monitor or off to the side so that the response key (see page 6) can be easily reached. The chair in which the subject sits should not be of the swivel, reclining or lounging variety; a standard desk chair would be optimal. In addition, the computer area should be clear of distractions and excessive equipment.

The room itself should be dimly lit, free of any source of light creating a glare. The room should be quiet and external noise should be minimal. If available, low level "white noise" can be present so as to block out any external and intermittent distracting activity.

As with any testing protocol, rapport should be established with the subject and any situational anxiety or oppositionality should be handled prior to the test session. Many intervening variables can affect the reliability and validity of the obtained results, as is true for many other psychometric test batteries. It is important to consider who will be administering the test. An individual who is familiar with attention deficits and how to interact with a person who presents with these deficits. A technician with no such knowledge will ineffectively instruct the subject about the test. A person only with computer knowledge would not be an effective technician -- the subject should remain the focus and not the hardware or software.

STARTING THE PROGRAM

If you have not already done so, check all computer connections and slot placements, specifically the clock, printer and the disk drives.

The proper disk placements are crucial. The CPT operating disk, i.e. the disk you received from the developers, must be in drive 1. A DOS 3.3 (see Apple reference manual(s) initialized disk to be used for data storage should be placed in drive 2. Obviously, make sure the drive doors are closed.

There are two ways the CPT can be accessed. One is through a warm-boot, the other is a cold-boot (see Apple reference manual). The cold-boot will result in a white band appearing at the top of the screen with the letters "CPT" superimposed. This will soon be replaced by another display with instructions for viewing an introduction text which parallels what is presented here. This section may be bypassed by pressing the [ESC] key. Scanning through the introduction is done by the ↓ and ↑ key.

Another way to access the CPT program is through the warm-boot. If this warm-boot method is used, it will result in further options being displayed on the screen after the white band described above is replaced.

Both methods, once completed, will bring the psychometrist to the MAIN MENU option display. This consists of four choices:

92

1) PRESENTATION [P]
2) SCORING [S]
3) REACTION TIMES [R]
4) EXIT PROGRAM [E]

Figure 1 displays the overall program configuration for each of these options and every section within them. This will provide the psychometrist with a reference to follow. The EXIT ([E]) option takes the psychometrist out of the CPT program; the program can be entered again at the same point by typing in RUN.

Presenting the test

If [P] is pressed, the major routine in control of presenting CPT will be loaded. This routine has two (2) subroutines, one which allows the psychometrist to enter subject demographics and test parameters, the other which accesses the main CPT program. In order to access this presentation program, the psychometrist must supply pertinent information through the demographic subroutine.

Demographic spreadsheet (option [1])

Once this portion has been loaded, the psychometrist will be presented with spaces into which information must be entered. The cursor will appear as a blinking square. Entry of information will require strict syntax. Table 1 will provide information on how this should be entered. This card indicates valid responses and examples. Only clarifications of these will follow:

NAME, ID#: If this information is not important, input of these two items can be skipped (↓ , [RETURN] or [TAB]).

SUBJ#: This must be entered in order for the data to be stored.[1]

AGE: If a child is below the age of 10, his/her age should follow a 0. (e.g. 07, "leading zero"). [see footnote -- *]

D.O.B.: Note syntax
1) Number day (e.g., 07, 31);
2) First three letters of month (e.g., NOV);
3) Last two digits of year (e.g., 84).*

SESSION: This is designed to document repeated testing. Leading zeros must be entered.[1]

GROUP: This item will set the testing inter-stimulus interval which has been derived from normative data. Ages on which this test has been normed are 7 through 12. These chronological ages are appropriate for input with leading zeros, if necessary. If a subject does not fit within this range, then either 06 or 13 are allowed. Refer to Table 2 for the resulting rates. The psychometrist may want to assign a higher "group age" to a child whose test date is within three (3) months of his/her next birthday.*

93

ADAPTIVE: If the adaptive version is selected, once the test begins, the starting rate will adjust according to the subject's performance. In particular, the stimulus presentation rate will increase if a correct response is made and decrease if any error is made. If the non-adaptive version is selected, the starting rate will be maintained throughout the test. We recommend the non-adaptive version.*

SERIES: Stored in memory are ten (10) one-hundred letter blocks (1-9 and A). These ten blocks have been put together to create ten different series. This feature was added to increase the flexibility of available letter sequence. Table 3 indicates these series.*

BLOCKS: The test can be given with up to ten blocks (1000 letters). Any number below can be selected. This, in conjunction with which series is selected determines which blocks will be utilized. It is recommended that the subject be given at least five (5) blocks (we prefer seven (7)) to measure in a reliable and valid manner the underlying characteristics for which the test was developed.*

START: This rate should have already been entered by the computer once the GROUP value was entered. However, this can be overwritten similar to any other supplied value within the Demographic Spreadsheet.*

MIN, MAX: These minimum and maximum rates are the limits of the interstimulus interval beyond which other perceptual variables are felt to effect the performance. These can also be changed if desired. However, if the non-adaptive version has been selected, these rates will be equal to each other. Limits of these values are .200 seconds (minimum) and 2.550 seconds (maximum).

If, after entering all necessary and desired information, the psychometrist wishes to change something, pressing [N] in response to the question:

"Information OK as shown?"

will return the cursor to the beginning of the spreadsheet. Use of the special function keys indicated on the Demographic help card will enable the psychometrist to move to the appropriate item for correction or change. Make sure that other items desired stay intact after this change (a warning message will appear if a correction is desired).

Once this information is accurate, [Y] in response to the same question the psychometrist will return to the local menu for the next selection (refer to Figure 1).

---

[1]The file name by which data will be stored and retrieved automatically by the listing:

1) The letters "CPT";
2) Four digit and/or letter subject number;
3) Session number.

EXAMPLE: CPTA20902

FOOTNOTE:

*Any invalid response will cause an error message to flash towards the bottom of the screen. Refer to the help card. Once the error has been located, the key, when pressed, will reinsert the cursor at the beginning of the item where the error occurred.

Color/letter grid (option [2])

As noted in Figure 1, this option must be selected in order for the parameters to be set. This grid is accessed by pressing [2] in response to the option menu on the screen. The psychometrist will see eight letters of different colors on the screen:

1) Dark green "U"
2) Purple "I"
3) Orange "A"
4) White "S"
5) Magenta "C"
6) Brown "L"
7) Yellow "E"
8) Dark blue "T"

If these colors do not appear as the correct color, the video monitor adjustment is essential before continuing.

At this point, the accuracy with which the subject can identify letters and colors can be determined by asking the subject to name each letter and its corresponding color. Some room should be given for personal interpretation (e.g., "red" for "magenta") but if there is strong evidence of color blindness or non-recognition of letters, the test should be stopped (as of yet, there is not a non-color version available). If the subject has no such deficits, the color version should be selected by pressing [Y]. Two target options will appear:

1) White "S"
2) Orange "S"

This refers to the color of the first letter of the target. The White 'S' (with a Blue "T") is preferred due to its clarity and display consistency on all monitors (whereas the color "orange" appears to vary in shade on different monitors). Table 4 presents the two different versions. Once a test version has been chosen, the computer will present the target letters and corresponding colors on the screen.

At this time, explanation of the test should begin and continue throughout the following sections when appropriate (refer to Testing Session Example on page 16). Once this has been done, the press of any key will return to the option menu. The psychometrist may either return to an already completed step (for further changes or clarification) — option [1] or [2] or leave the Presentation program (option [6]).

10 Letter Introduction (option [3])

This step is very useful to the subject due to its brevity and preparatory function. If selected, it will provide a ten (10) letter sequence at the rate with which the subject will see the letters during the Main Sequence. The sequence for the White "S" version is as follows:

1) Green "I"
2) Yellow "C"
3) Brown "E"
4) White "S"
5) Blue "T"
6) Purple "H"
7) Orange "O"
8) White "S"
9) Green "T"
10) Magenta "A"

If the countdown and the inter-stimulus interval is much different in rate from what was set with the Demographics, this indicates that the clock has not been properly modified for a 10 msec. interrupt. Note also that within this sequence there is a "true" target (letters #4 and #5) and a "false" target (#8 and #9) (see Explanation of Results for complete explanation (page 10)). The subject should perform perfectly during this 10 letter sequence, even if it requires multiple trials. This enables the psychometrist to be sure that the subject understands the task and is responding accurately.

NOTE: The response key is the "CLOSED APPLE" [ ] key. No other key on the keyboard will register any response.

100 Letter Practice (option [4])

Although optional, this sequence is highly recommended because it will indicate to the psychometrician if the subject understands the task. Clinical judgment will be necessary, since errors may either be part of a comprehension problem or a manifestation of Attention Deficit Disorder symptoms. It is esential that the psychometrist stay with the subject to answer any questions. If at any point it is necessary to stop the practice sequence due to an obvious error in comprehension, he/she may press [CTRL] [S]. Once clarification has been given, the sequence may be resumed ([R]). Because this feature is available during the Main Sequence (option [5]) and can be easily accessed by th subjects, it will be essential to hide this key sequence from the subject so he/she cannot stop the test.

The 100 letter block chosen for the practice sequence is Block #7. Note in Table 3 that this block will not appear in the main test sequence unless the test's length is over seven blocks.

Once the practice sequence is completed, the data will automatically be analyzed but not stored on disk. The indication that analysis is proceeding properly is a thick solid line building across the screen until the printing options (see Displaying Data, page 8) appear. Once the results have been scanned, basic encouragement should be given to prepare the subject for the Main Sequence.

Main Sequence -- testing session -- (option [5])

After a few preparatory steps, the countdown will appear followed by the letter sequence. During the test, the psychometrist need not stay in the room with the subject unless the subject is unduly anxious or it has been determined that his/her presence will not distract the subject. REMEMBER, THE RESPONSE KEY IS THE "CLOSED APPLE" [ ] KEY. NO OTHER KEY WILL REGISTER A RESPONSE.

CODE AND BLOCK DEVELOPMENT

This section explains the development of the ten 100 letter sequences. This will aid in understanding the scoring procedures.

To make this explanation easier to understand, it will be assumed that the target sequence is White "S", Blue "T". Other parameters are as follows:

1) Ten (10) targets (true target) appear in each block. This comprises twenty (20) letters;

2) Twenty (20) letter pairs appear in each block that has three of the four characteristics of the target. that is, these are "false" targets (target-like) pairs):
   a) White "S", Blue "non-T"
   b) White "S", non-blue "T"
   c) White "non-S", Blue "T"
   d) non-White "S", Blue "T"

   This comprises forty (40) letters. The fourth erroneous characteristic may be that of another target characteristic (e.g., "non-T" = "S"), but due to randomization this appears three or four times within 1000 letters;

3) The remaining forty (100-(20+40)=40) letters ("fill-in" letters) are dispersed throughout each block. These letters have none of the four characteristics of the "true" target (i.e., colors: Blue, White; letters: "S", "T");

4) No "true" target or "false" target pair will follow the same category. For example, these will never occur:
   a) White "S", Blue "T" followed by White "S", Blue "T"
   b) White "non-S", Blue "T" followed by White "non-S", Blue "T";

5) Each "true" target or "false" target pair will be followed by at least one fill-in letter, sometimes two and rarely three;

6) Linking blocks together to create the sequences will not result in any additional "true" target or "false" target pair;

7) The first letter of any block which is first in the test sequence will not be part of either a "true" target or "false" target pair. The "true" target or "false" target pair will not appear at th end of any block which is last in the test.

   See Table 10 for complete letter/color sequences for each block.

DISPLAYING DATA

Two types of data displays are available. These displays are accessible only after analysis has been completed. It is important to note that results of analyses are not stored on disk, only raw data. Once the psychometrist leaves the analysis routine, any clinically useful and composite results will have to be generated again through the analysis procedures (Scoring and Reaction Times). Two display modes are possible:

[1] This will display the analyzed data on the video monitor in a format which is readable (only if the 80-column board in the computer is in place). When the display has scrolled through [RETURN] will enable the psychometrist to select another display option. If the psychometrist wishes to temporarily stop the scrolling [CTRL] [S] will do so. Pressing any key will resume the scrolling once it has been temporarily stopped.

[2] (a) This option will output the analyzed results onto the printer. Once this option is selected the psychometrist will be instructed to prepare for printing and [RETURN] will activate the printer. Following printing, another [RETURN] will return the psychometrist to the local menu.

(b) An option [T] will be available to append a one page explanation (Text) of all possible errors used within the Scoring output will be readily decodable.
(The Scoring results generated and appended Text will be discussed in the section entitled Explanation of results (page 10)).

[3] This allows the psychometrist to return to the Main Menu for another selection.

ANALYSIS OF DATA

Once the presentation phase of the CPT has been completed, the data will automatically and immediately be stored under the name of CPTTEMP.DATA on the CPT system disk (drive #1). This is to insure that the immediate data would not be lost in the event of a missing or faulty data disk in drive #2 (see footnote). Once the temporary data has ben stored, it will be copied over to the data disk in drive #2 under the concatenated filename with the subject number and session number preceded by 'CPT'.

An existing or newly created data file which contains the time at which a response occurs can be scored in two different ways. Actual error classifications are not done until the analysis is completed. Specifically, analysis is completed "off-line" and not while the test sequence is being presented.

Scoring

After successfully completing this, the psychometrist may return to the Main Menu where data could be scored [S]. If a CPT test was given prior to accessing the Scoring subroutine, selection of option [0] will score this current file. If no test resides in current memory or the psychometrist wishes to score a different file, selection of option [1] will ask the psychometrist to provide a valid and existing filename.

Once this is provided, the computer will indicate which series (block sequence) was used. The following phase of analysis will take between three (3) to fifteen (15) seconds to complete, depending on the number of blocks which were used for the test session. This phase is called "number crunching" and will be indicated by the solid bright line building successive rows until all data has been analyzed.

Reaction Times

This is a supplementary analysis of the reaction times with which the subject responded and a running list fo correct responses, types of errors and when they occurred. The procedure to obtain an output uses the same procedure as that for Scoring. These results can be useful for purposes of response pattern analysis and a basic physiological measure of reaction time. Again, specific explanations will be given in Explanation of Results (page 10).

Footnote: If a disk in drive #2 is missing, the drive door is open or a duplicate filename is found, the psychometrist will see an error message and be given options to remedy the problem. If it is the case of a duplicate filename, the program will return to the subject number and session number in the Demographic spreadsheet to correct this duplication.

EXPLANATION OF RESULTS

Scoring and Text

Table 5 presents an example of the error classification printout produced by the Scoring routine. The demographic data inputted during the preparation of presenting the test is also printed and is part of the subject's permanent computer data file. Table 6 presents the Text printout which may be appended to the error classification printout only if option [2] of DISPLAYING DATA (page 8) is selected. The following section will explain the types of responses recorded. The section CODE AND BLOCK DEVELOPMENT should be used for reference to accompany this explanation. In addition, a few terms need to be defined:

A) 'True' target: This is a two letter sequence presented in appropriate colors (i.e. White 'S', Blue 'T') which if responded to would preclude an Omission.

B) 'False' target: A two letter sequence which has only three (3) characteristics of the 'true" target (see CODE AND BLOCK DEVELOPMENT, page 7).

C) window: CPT scoring is based on a letter interval paradigm. Each occurrence of a letter marks the start of another time interval which is determined by the setting of rate (start, win, max) parameters. The rate is also an inter-stimulus interval. A 'true' or a 'false' target response is scored in relation to the number of intervals that have been presented since the onset of the second letter of the 'true' or 'false' target. The operational definition of an open (or within) window response is a response occurring within two intervals following the onset of the second letter of a 'true' or 'false' target. A closed (or outside) window response is a response occurring after two intervals following the onset of the second letter of a 'true' or 'false' target, but before the first letter of another 'true' or 'false' target has been presented.

The error classification is presentd in three sections. The first section is a composite profile of the subject's overall performance:

RESPONSES: This is the number of key presses recorded while the subject was responding correctly or incorrectly. If for each block, the subject responded only to a correct target, the number "10" would appear.

CORRECT: This represents the number of 'true' targets to which the subject responded correctly. Since there are exactly ten (10) targets per block, a percentage value can be derived.

ERROR: This section presents the number of errors, regardless of the type which the subject committed. This section is further broken down in the second section of the printout.

The second section classifies errors according to the major categories -- Omissions and Commissions.

OMISSION: These numbers and percentages will obviously be the arithmetic complement of the number of correct responses. An operational definition of an Omission is as follows:

> **A failure to press the 'closed apple' [[illegible]] key during the open window rsponse interval (within two intervals).**

LATE OMISSION: This second type of Omission error is scored if the subject responds but during a closed window. The number of these errors possible is dependent on which blocks are being presented and how many "fill-in" letters occur between 'true' and 'false' targets. Table 7 presents the maximum number of the late omission responses possible for each block. Refer to COMMISSION TYPE III-A for a differential classification of these errors.

COMMISSION TYPE I: The error is registered when an open window response occurs in reaction to any of the four 'false' targets (Refer to CODE AND BLOCK DEVELOPMENT). The third section of this printout expands these responses to the frequency with which each 'false' target response occurred.

COMMISSION TYPE II: This Commission error category is similar in definition to the LATE OMISSION (applied to closed window responses) but indicates the frequency of a 'false' target rsponse. Table 7 presents the maximum number of these commission errors possible. Differential classificiations are possible of these errors. (Refer to COMMISSION TYPE III-A).

COMMISSION TYPE III: This category can provide important information as to whether the test can be considered valid. In formulating this error classification the goal was to provide a categorization of errors specific to each behavioral response. The following categories can provide crucial information regarding the subject's comprehension of the test, anticipatory tendencies and over-responsivity to the stimulus separate from the COMMISSION TYPE I and II errors:

COMMISSION TYPE III-A: This error category includes all responses made to the first letter of either a 'true' or 'false' target before the second letter appears. These are classified an Anticipatory Response errors. These responses may ocur in individuals who require more time than is given within the response window. Therefore, these responses may be a reflection of the number of "fill-in" letters before the anticipatory responses and/or the neurological intactness of the subject.

These errors may more appropriately be LATE OMISSION or COMMISSION TYPE II. If these errors occur frequently, the Reaction Time scoring routine can provide additional information regarding the feasibility of an anticipatory response in relation to a minimum physiological speed with which to decode a stimuli. Refer to Reaction Time from further explanation (page 13).

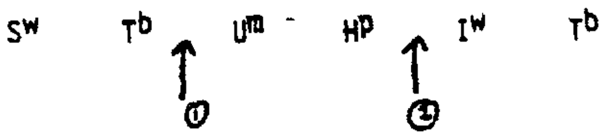
<u>COMMISSION TYPE III-B</u>: These errors will be recorded if a response is made in a subsequent letter interval although the subject responded in the preceding interval. These are called Additional Response errors. These errors can be generated by responding either to a 'true' or 'false' target. These types of errors can be subdivided:

Open window: An additional response (as described above) during the allowed two letter interval which is operationally used as the active response window. For example, with the arrows indicating a response and the superscript letter indicating the color of the letter:

The first rsponse would be recorded as a correct response to a target. The second response would be recorded as an <u>open window</u>.

COMMISSION TYPE III-B, since the latter response occurred before the onset of the purple 'H' which marks the close of the target window. If the 'S' was of a different color, the second response would stay the same, but the first response would be considered as a COMMISSION TYPE III-D (See CODE AND BLOCK DEVELOPMENT).

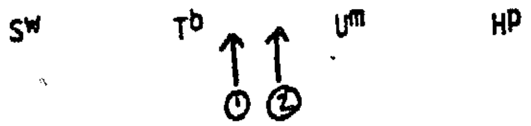
Closed window: This definition logically follows the previous error. An example should suffice:



The second response is after the active response window has closed but before the onset of a letter which is part of a 'false' target (White non-S, Blue 'T').

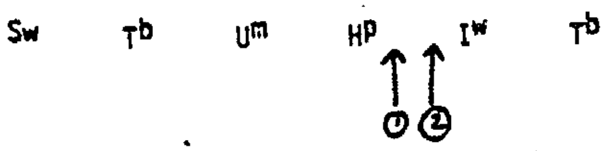
COMMISSION TYPE III-C: This error will be recorded when a subject responds two or more times within the same letter interval. This is referred to as a Multiple Response error:

Open window: This category will be chosen for a response such as this:



would generate a correct response and a COMMISSION TYPE III-C open window error.

Closed window: This response style:



would be described as one LATE OMISSION and one COMMISSION TYPE III-C closed window response. If the target was instead, a 'false' target, the first response would be considered a COMMISSION TYPE II error; the error category for the second response would remain the same.

If either of the two previous categories (COMMISSION TYPE II and III) occur with a relatively high frequency, one may speculate a variety of task inappropriate behaviors ranging from poor comprehension to oppositionality to extreme impulsivity. Although these errors alone should not invalidate the test, consideration should be given to the degree of off-task behavior alters the maningfulness of the test and the behavioral problems it is designed to measure.

Reaction Times

At first glance, the Reaction Time printout looks incomprehensible. However, if one remembers that the CPT scoring is based on a letter interval response paradigm, appropriate definitions begin to fall in place.

Any display of the reaction time data will include the demographic data. The columns numbered 0 through 16 will appear before the first response of each block administered. These numbers correspond to the response categories in the Scoring prinout (see Table 5 and 9). For each response made, an asterisk will appear below one category (0-16) preceded by five values which can be used to pinpoint when the response was made. STIM denotes the nth letter in which the response occurred.

All other categories (EVENT, cTIME, RATE, tTIME) indicates values read from THE CLOCK". Each of the numbers signify the number of 10 milliseconds time intervals for that category. For simplicity, since the test begins at time = 0, multiplying the number by 10 indicates the time that has elapsed. More specific explanations for these categories will make the scoring routine more clear:

RATE: This indicates the length of the letter interval in which the response occurred (40 = 400 milliseconds). This value, if using the non-adaptive version of the test, should be the same as the starting, minimum and maximum rates set in the demographic spreadsheet and should remain the same through the entire test.

cTIME: This value indicates the time by which the interval ends for the letter in which a response occurred. As shown in Table 8, this person responded for the first time during the sixth letter whose time interval closes at cTIME = 240 (2400 milliseconds from the beginning of the test).

EVENT: This is simple the time when the actual response occurred. Therefore, knowing cTIME and STIM, EVENT identifies exactly where within the STIM interval the response was made.

tTIME: This value is the length of time that has elapsed since the onset of the first letter of either a 'true' and 'false' target given that both of the letters of either target has been presented. More specifically, if tTIME is positive, this indicates that a response was made after the second leetter of either a 'true' or 'false' target. If, however, tTIME is negative, this mans that a response was made before the onset of the second letter (i.e., Anticipatory). The value indicates the length of time still to elapse before the onset of the second letter. Any positive value over 120 (1200 milliseconds) indicates a Late Response. Obviously, to obtain a reaction time, one needs to subtract 40 from tTIME to determine the elapsed time since onset of the second letter of either 'true' or 'false' target. The formula for Reaction Time is therefore:

RT = 9tTIME * 10) - (RATE * 10)

104

Illustrations of these examples will clarify the Reaction Time printout. See Table 8 (Response A).

A --- stimulus#

stimulus#	4	5	6	7	8
$letter^{color}$	D^{m}	S^{w}	T^{b}	U^{o}	Y^{w}
Interval close time(cTIME)		160	200	240	
Response (STIM) = 6)					
Response time (EVENT)			↑ 224		

tTIME = 64

In response to the Blue 'T' of a 'true' target, the subject responded 240 milliseconds into the Blue 'T' interval (RT = 640 - 400). An asterisk appears in column 0 of STIM - row 6.

B --- stimulus#

stimulus#	311	312	313	314	315	316
$letter^{color}$	P^{p}	S^{w}	T^{b}	E^{o}	H^{g}	U^{w}
Interval close time(cTIME)		12440	12480	12520	12560	12600
Response (STIM)						
Response time (EVENT)					↑ 12570	

tTIME = 131 (RT = 1310 - 400)

C --- stimulus#

stimulus#	265	266	267	268	269
$letter^{color}$	S^{w}	T^{b}	L^{y}	S^{g}	T^{b}
Interval close time(cTIME)		10600	10640	10680	10720
Response (STIM)					
Response Time (EVENT)				↑ 10704	

tTIME = -16 (Anticipatory response Type 9)

In this case, the usual RT formula cannot be used (whenever tTIME is negative). Instead the length of time that must elapse until the onset of the second ltter is | tTIME | * 10 (absolute value of tTIME * 10 = 160 milliseconds).

NOTE: This is a case where a differential classification may be speculated. The sequence before the response included a 'true' target followed by one "fill-in" letter. It is possible that the subject required more time in which to respond correctly than the sequence allowed. This re-classification should be made on the basis of the responses the subject gave throughout the entire test. If anticipatory errors occur frequently following a target with only one "fill-in" letter, this indicates that the subject needs more time in which to respond (>RATE * 2 milliseconds). If such a response follows a 'false' target, this may suggest immediate or 'delayed' impulsivity.

105

TESTING SESSION EXAMPLE

The following describes a sample testing session. This may be used as a model by which to guide oneself through the test instructions given to the subject and appropriate responses to screen displays. U will signify the psychometrist dialogue; S will signify the subject dialogue. D will signify the screen display and I will signify information typed in on th keyboard.

Prior to the subject entering the room, the Presentation program has been loaded [P] and subsequently the Demographic Spreadsheet [1]. Demographic information has been entered and verified as correct. The parameters set for data storage and test administration are:

	Display	Input
1)	Subject#:	A101
2)	Sex:	M
3)	Ses.ion:	01
4)	Group:	10
5)	Adaptive:	N
6)	Series:	01
7)	Blocks:	07
8)	Start:	already set at .600 due to input as a 10 year old male
9)	Min,Max:	already set at .600 due to input as the non-adaptive rate version.

D: "Information OK as shown?"
I(U): [Y]
D: ** CPT PRESENTATION **
[1] Enter DEMOGRAPHICS
[2] Color/letter GRID
[3] EXIT to main menu

Perform which function? [X]

I(U): [2]
D: Eight colored letter appear in a 3X3 grid
"Color version (Y or N)?" [X]

The subject is brought in. If the psychometrist does not know the subject, a few minutes should be taken to establish rapport, which should include obtaining information about his/her familiarity with computers. The subject should be comfortably sitting in a chair in front of the computer video monitor. Equipment rearrangement may be necessary.

U: "This is a game to see how well you can pay attention to what will appear on the screen. But first, I'd like to find out if you can see letters and colors okay.. this is a green 'U', right?" (point to upper left letter). "What's this?" (point to upper middle letter; repeat this procedure as necessary*).
I(U): [Y] (If subject identifies the letters and colors adequately)
D: Same eight letter grid with:
[W] White-S
[O] Orange-S
Which option: [X]

I(U): [W]
D: A White 'S', Blue 'T' appears in the middle of the screen
"Press ANY KEY to RETURN to menu" [X]

U:** "I'd like to explain exactly what you have to do for this game. You will have a chance to practice, so don't worry if you're not sure about what it is you're supposed to do. Feel free to ask any questions about the game itself. It's very important that you ask questions if you have any, because it wouldn't be fair to you if I'm asking you to do something that you're confused about. Many children have questions about this game, so please don't feel that you're dumb if you don't understand something I said, okay?"
S: "What's this thing?" (points to printer, if in sight)
U: "It's a machine that types things out, but you won't be using it. The only thing that you have to use is one key on this keyboard while you're looking at this screen. Once we get started, you'll be seeing letters flashing on the screen one right after another. And the letters will be in different colors, too. So, one letter will flash on the screen (a hand motion to demonstrate a 'flash' may be helpful), then that letter will go off but another will come on right after that."

---

* The psychometrist may correct the subject after he/she has completed the identification phase. A determination will have to be made whether or not these errors warrant discontinuation of the test.
** This, of course, is an exampl of a smooth introduction. The psychometrist should feel free to deter from this presented method according to the needs of the subject.

"Sometimes you may see a White 'S' (point to the screen) and then right after that a Blue 'T' may appear. If you see these two letters appearing one right after the other, in these colors and in this order (frequently refer to the target appearing on the screen), prss this key (point to the 'closed apple' [ ] key). It's important that you press only this key because the computer won't know that you'r paying attention if you press other keys."
S: "What if I see an 'S' and a 'T' in different colors?"
U:* "That's a very good question! The only time you want to press this key is if the 'S' is a White 'S' and the 'T' is a Blue 'T'. The computer will try to trick you by showing you all different combinations of letters and colors which are very close to the White 'S', Blue 'T', but pressing the key for anything other than the White 'S', Blue 'T' would be a mistake. If you do make a mistake, don't get mad or frustrated with yourself -- everyone makes some mistakes. Just do the best you can."
S: "Do I press this button after the White 'S'?
U:* "No, it's important to press this button (point) after the Blue 'T' only if the White 'S' comes before it. Just because you see a White 'S' doesn't mean you should press this button; you should wait to see if the Blue 'T' comes next because the computer may try to trick you like I mentioned before. Why don't you practice a bit to show me that you understand what you're supposed to do."
I(U): (Any key pressed will advance the program)
D: (The CPT Presentation menu has appeared again with three additional options -- [3], [4], and [5])
I(U): [3] (10 letter INTRODUCTION)
D: "Loading Display Routine" (3 seconds)
(New Screen) "Press ANY KEY to BEGIN Series"
I(S): Subject presses a key and sequence starts.
U: "Always let me set the game up; you'll have a chance to play in a second" (The sequence ends).
D: "Press ANY KEY to RETURN to menu"
I(U): (Any key is pressed followed by option [3])
U: "I'm going to show you 10 letters one right after another. What are you supposed to do when you see the different colored letters?"
S: "I should press this key (points to the 'closed apple [ ] key) whenever I see a White 'S' and a Blue 'T'."

---

* Make sure to include these explanations even if the subject does not ask the above question. This is crucial to the subject understanding the test.

U: "That's right. Press this key if you see a Blue 'T' only when a White 'S' came before it. Good! In a few seconds, you'll see a countdown with numbers which tells you to get ready. Then some letters will start flashing. There may be a White 'S', Blue 'T' in there; if so, press right after you see the Blue 'T'," (Make sure the subject is prepared; have subject put finger right above the 'closed apple' key as if ready to press)
I(U): (Any key is pressed)
D: (Countdown starts from the number 5 and continues at the rate at which the psychometrist requested.
I(U): Subject presses both for the 'true' and 'false' target. The 'false' target response is followed by "Oops".
U: "The computer caught you! See, that's how the computer will try to trick you. Try your best to pay close attention.* Okay, let's try it again (Return to 10 letter introduction option and begin sequence).
I(U): The subject responds to the 'true' target correctly and appropriately inhibits a response to the 'false' target.
U: Excellent! You did a perfect job! Now, let's have some more practice. (Return to menu and choose option [4] -- 100 Letter Practice)
D: "Loading code for practice sequence" followed by
"Press ANY KEY to BEGIN Practice"
U: "You'll see a few more letters, now. Ready?
I(U): (A key press begins the sequence)
U: (Observe the subject's responses to determine comprehension and attitude toward the test. Discuss blatant errors such as multiple or random responses whether or not they're associated with the stimuli and return to previous steps where appropriate) "Okay, let's see how you did."
D: "Analyzing Practice Results" (The practice sequence will automatically be analyzed; consecutive solid lines will be forming).

[1] Display Results on 80 col. MONITOR
[2] Output Results to PRINTER
[3] EXIT to local menu

---

* If it is evident that the subject is waiting a long time to respond correctly or if he/she hesitates to press because one letter has gone by after the Blue 'T' (i.e. the window is still open), encourage him/her to respond as soon as he/she can.

I(U): [2] (It is best to have a permanent copy of the practice session -- practice results are not stored on disk). Printout shows nothing indicative of non-compliance or poor comprehension. Therefore, continue
U: "Okay, now we're ready for the real game."
I(U): Return to option menu [RETURN] then [3] for EXIT to Local menu and select [5] -- Main Sequence.
D: "Loading Series #1"
"Press ANY KEY to BEGIN sequence"
U: I'm going to leave the room (or "I'll be back here") so I won't disturb you. Do the best you can and remember, White S, Blue T (sequence begins).
D: (Subject leaves following completion of the test sequence)
"**PLEASE DO NOT DISTURB**"
"STORING CPTTEMP.DATA on sys disk"
"STORING CCPTA10101 on data disk"*
(New Screen)
"FILE:CPTA10101 ALREADY EXISTS"
"for SUBJECT #A101"
SESSION #01
"Program WILL NOT overwrite"
"Enter RETURN to correct demographics"
I(U): [RETURN]
D: "Loading demographic spreadsheet" (program will return cursor to SUBJECT # and SESSION to allow for correction).
I(U): (Changes SUBJECT # to A102 and [TAB] through SESSION)
D: "Information OK as shown?"
I(U): [Y]
Another possible error which may occur at this point is:
D: "Storing CPTA10101 on data disk"
(New Screen)
"I/O ERROR"
"CHECK:"
"1) Is the CPT-SYSTEM Disk in D1?"
"2) Is the DATA Disk in D2 INITIALIZED?"
"3) Are the Drive Doors SHUT?"
"Please correct the problem"
"Then press RETURN to continue"
I(U): (Drive #2 door was not shut) [RETURN]
D: "Storing CPTA10101 on Data Disk

---

* The following steps will not occur unless there has been an error in how the test results will be stored. Refer to the step marked by the check mark ( ).

( ) (New Screen)
"***CPT PRESENTATION***"
[1] Enter DEMOGRAPHICS
[2] EXIT to main menu
"Perform which function?" [X]
I(U): [2] (If the psychometrist wishes to administer another test, [1] will lead back to the point where the TESTING SESSION EXAMPLE began).
D: (New screen) -- "CPT Main Menu"
I(U): [S] -- Scoring
D: "Loading Scoring Routine"
"***CPT SCORING***"
[0] SCORE Current File CPTAA10201
[1] LOAD New file
[2] EXIT to main menu
"Perform which function? [X]
I(U): [0]
D: "Loading series #1"
"Busy number crunching (solid line extends across screen)"
(New Screen)
"[1] Display Results on 80 col. MONITOR
[2] Output Results to PRINTER
[3] EXIT to local menu"
"Perform which function?" [X]
I(U): [2], [RETURN] (Printer prints out results similar to Table 5)
D: "[T] Append explanation TEXT
[R] RETURN to menu"
"Perform which function? [X]
I(U): [T] (Printer will print out Table 6)

After TEXT is printed, the computer will return to the printing options. [3] will return to the Scoring options; a [2] selection will return to the Main Menu, where the psychometrist can set up for another Presentation, Scoring, Reaction Time or EXIT. If desired, typing in RUN will return the psychometrist to the point where the program was exited.

\\\

REFERENCES AND BIBLIOGRAPHY

Garfinkel BD, Klee SH (1983) A computerized assessment battery for attention deficits. The Psychiatric Hospital 14(3):163-166.

Garfinkel BD (1984) Neuroendocrine and cognitive responses to amphetamine in adolescents with a history of ADD. In: Attention Deficit Disorder: Diagnostic, Cognitive and Therapeutic Understanding. (Ed.) L. Bloomingdale, Spectrum Publications Inc.: New York.

Klee SH, Garfinkel BD (1983) The computerized continuous performance task: A new measure of inattention. Journal of Abnormal Child Psychology 11(4):487-496.

Shapiro SK, Garfinkel BD (submitted, 1985) The occurrences of behavior disorders in children: The interdependence of Attention Deficit Disorder and Conduct Disorder.

#164, SS-SS8

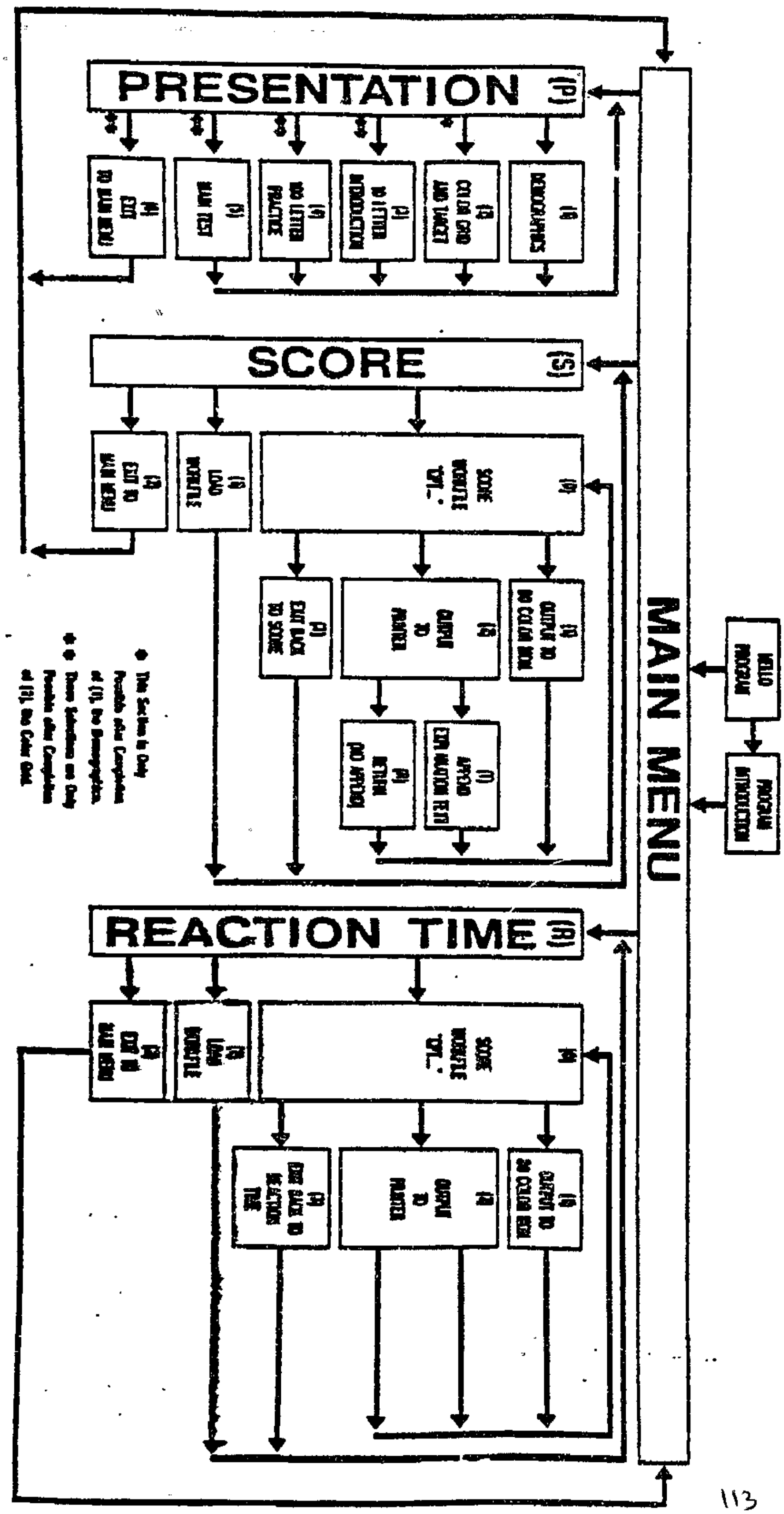


FIGURE 1

TABLE I
DEMOGRAPHIC HELP CARD

Field	Valid Responses	Number of Digits	Examples; Comments	Notes
Name	Letters, numbers, punctuation	24	Smith, John Henry	
ID#	Letters, numbers, punctuation	7-1	D134729-2; Hospital or Clinic ID	
Subj#	Letters, numbers	4	A209; Subject number	1
Age	00,01,02.........99	2	30; Biological Age (years)	
Sex	M or F	1	M	
D.O.B.	01 to 31--3 letters--00 to 99	2-3-2	30-AUG-54; Date of Birth	2
Date	01 to 31--3 letters--00 to 99	2-3-2	20-SEP-84; Test Date	2
Session	00,01,02.........99	2	02	1
Group	06,07,08.........13	2	13; Group reflects the test age	3,4
Adaptive	Y or N	1	Y=Adaptive Rate; N=Non-adaptive Rate	5
Series	01,02,03.........10	2	Each series is a fixed, random combination of ten blocks.	
Blocks	01,02,03.........10	2	Each block contains 100 letters. This entry determines the length of the test.	
Start	0.200 to 2.550	4	Rate (in seconds) at which letters appear	4,5
Min	0.200 to 2.550	4	Fastest rate attainable	4,5
Max	0.200 to 2.550	4	Slowest rate attainable	4,5

Special Function keys:

- Right Arrow: Moves the cursor right one character position.
- Left Arrow: Moves the cursor left one character position.
- Down Arrow: Moves the cursor to the first character position of the next field.
- Up Arrow: Moves the cursor to the first character of the previous field.
- Return: Same as Down Arrow.
- Tab: Same as Down Arrow.

Notes:

1) The subject and session fields are combined to form the file name. Using the examples from each field respectively, the created file name would be: CPTA20902.
2) **Do not use** 8-30-54 for 30-AUG-54. The computer will not accept it.
3) The **group age** (test age) corresponds to chronological age in years; round to higher age when test date is within three months of next birthday.
4) Age appropriate rates are automatically inserted according to the value of the **Group** (Refer to **Table 2**: Default Presentation Rates).
5) If the non-adaptive selection is chosen (i.e., N) then **Min** and **Max** will automatically be set to the same rate as **Start**.

114

TABLE 2

DEFAULT PRESENTATION RATES*

Group (years)	Male (seconds)	Female (seconds)
< 7	.990	.720
7	.990	.720
8	.700	.640
9	.670	.640
10	.600	.640
11	.530	570
12	530	.520
13	.400	400
> 13	.400	400

* These will be the rates at which the letters are presented at the beginning (with the adaptive rate) or throughout the test. They can be changed if desired within the **Demographic Spreadsheet.**

TABLE 3

SEQUENCE OF BLOCKS FOR EACH SERIES

Series	Sequence of Blocks									
1	4	9	1	6	2	5	A	7	3	8
2	5	3	9	8	2	4	1	A	7	6
3	8	2	6	4	1	3	A	4	7	5
4	9	5	3	A	1	4	8	6	2	7
5	2	A	5	1	6	8	3	7	4	9
6	4	8	2	5	3	A	1	6	7	9
7	8	A	4	6	1	3	9	2	5	7
8	5	1	6	3	9	4	8	A	7	2
9	9	3	6	4	1	5	A	7	2	8
10	4	A	8	1	6	9	5	2	3	7

116

TABLE 4

COLOR USAGE FOR CPT VERSION

Color Name	Test Version* White S Blue T	Test Version* Orange S Blue T
Magenta	X	X
Dark Blue	X	X
Purple	X	X
Dark Green	X	X
Brown	X	
Orange	X	X
Pink		X
Green		X
Yellow	X	X
White	X	

* Colors used for the two target options available.
These steps were made to change the White S, Blue T version to the Orange S, Blue T:

1) Orange --------> White
2) Pink ----------> Orange
3) Green ---------> Brown

TABLE 5
Scoring printout

Name: Doe, John ID#: 11111111 Subject #: A101
Age: 25 Sex: M DOB: 21-FEB-59 Test Date: 04-FEB-85 Session #: 01
Group: 13 Adaptive: N Series#: 01 # of Blocks: 04 Color: W
Starting Rate: 0.400 Minimum Rate: 0.400 Maximum Rate: 0.400

Block	Responses #	Correct #	Correct %	Error #
1	13	8	80%	7
2	7	6	60%	5
3	12	9	90%	4
4	12	9	90%	3
Total	44	32	80%	19

Block	Errors						
	Omissions		Late Omissions	'Commission Type' Errors			
	#	%	#	TYPE I #	TYPE I %	TYPE II #	TYPE III #
1	2	20%	0	4	20%	0	1
2	4	40%	0	1	5%	0	0
3	1	10%	0	2	10%	0	1
4	0	0%	1	2	10%	0	0
Total	7	18%	1	9	11%	0	2

Block	Type I (expanded)								Type II (expanded)			
	A		B		C		D		A	B	C	D
	#	%	#	%	#	%	#	%	#	#	#	#
1	3	60%	0	0%	0	0%	1	20%	0	0	0	0
2	0	0%	1	10%	0	0%	0	0%	0	0	0	0
3	2	40%	0	0%	0	0%	0	0%	0	0	0	0
4	0	0%	2	20%	0	0%	0	0%	0	0	0	0
Total	5	25%	3	15%	0	0%	1	5%	0	0	0	0

Block	Type III (expanded)						
	A			B		C	
	1	2	3	1	2	1	2
1	0	0	0	1	0	0	0
2	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0
4	0	0	0	0	0	0	0
Total	0	0	1	1	0	0	0

118

TABLE 6
TEXT printout

TERMS USED

A) (TRUE) TARGET: White 'S' immediately followed by Blue 'T'

B) FALSE TARGET: Two letter sequence having only three (3) characteristics of true target (e.g., Green 'S', Blue 'T'; see TYPE I: A-D below)

C) WINDOW: Two-letter interval beginning from the onset of the second letter of a true or false target

1) OPEN WINDOW: less than two-letter intervals have passed

2) CLOSED WINDOW: Two letter intervals have passed

DEFINITIONS FOR ITEMS SCORED

--- CORRECT: Open window response to target

--- OMISSION errors: No response to target

--- LATE OMISSION errors: Closed window response to target (*)

--- COMMISSION TYPE I errors: Open window response to a false target (*)

A) White 'S', Blue non-'T'
B) White 'S', non-Blue 'T'
C) White non-'S', Blue 'T'
D) non-White 'S', Blue 'T'

--- COMMISSION TYPE II errors: Closed window response to a false target

A - D) Same as for 'TYPE I errors'

--- COMMISSION TYPE III errors: Additional errors

A) Response to first letter of a true or false target

1) White 'S'
2) White non-'S'
3) non-White 'S'

B) An additional response within a different letter-interval than an already recorded response

1) Open window
2) Closed window

C) An additional response within the same letter-interval as an already recorded response

1) Open window
2) Closed window

* The definition of these errors appear incorrectly on the computer generated printout

TABLE 7

MAXIMUM NUMBER OF CLOSED WINDOW RESPONSE ERRORS POSSIBLE*

Block	Late Omissions	Commission Type II
1	2	5
2	1	6
3	5	3
4	3	5
5	3	5
6	2	7
7	5	3
8	5	3
9	4	4
A	3	4

* Determined by random placement of "fill-in" letters occuring after a 'true' or 'false' target.

TABLE 8

Reaction Time printout

Name: Doe, John ID #: 11111111 Subject #: A101
Age: 25 Sex: M DOB: 21-FEB-59 Test Date: 04-FEB-85 Session #: 01
Group: 13 Adaptive: N Series #: 01 # of Blocks: 04 Color: W
Starting Rate: 0.400 Minimum Rate: 0.400 Maximum Rate: 0.400

STIM	EVENT	cTIME	RATE	tTIME	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
6	224	240	40	64	•																
12	467	480	40	67		•															
13	498	520	40	98						•											
21	811	840	40	51	•																
31	1230	1240	40	70	•																
38	1505	1520	40	65					•												
45	1785	1800	40	65	•																
51	2031	2040	40	71	•																
58	2303	2320	40	63	•																
67	2647	2680	40	87		•															
79	3152	3160	40	72	•																
86	3435	3440	40	75	•																
89	3552	3560	40	72		•															

STIM	EVENT	cTIME	RATE	tTIME	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
104	4125	4160	40	83			•														
146	5826	5840	40	66	•																
161	6434	6440	40	74	•																
172	6870	6880	40	70	•																
185	7391	7400	40	71	•																
191	7632	7640	40	72	•																
198	7911	7920	40	71	•																

STIM	EVENT	cTIME	RATE	tTIME	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
204	8150	8160	40	70	•																
213	8505	8520	40	65	•																
230	9190	9200	40	70	•																
235	9375	9400	40	95		•															
243	9711	9720	40	71		•															
247	9874	9880	40	74	•																
259	10353	10360	40	73	•																
268	10704	10720	40	-16										•							
272	10878	10880	40	78	•																
279	11135	11160	40	95	•																
285	11361	11400	40	81	•																
293	11716	11720	40	76	•																

STIM	EVENT	cTIME	RATE	tTIME	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
303	12118	12120	40	78	•																
315	12571	12600	40	131											•						
323	12912	12920	40	72			•														
336	13439	13440	40	79	•																
342	13677	13680	40	77	•																
348	13912	13920	40	72	•																
355	14191	14200	40	71	•																
362	14468	14480	40	68	•																
368	14716	14720	40	76	•																
375	14988	15000	40	68	•																
389	15555	15560	40	75	•																
400	15961	16000	40	81			•														

TABLE 9

ERROR CODES USED IN REACTION TIME PRINTOUT

Code*	Type of Error	Description
0	(CORRECT)	Punctual response to White S/Blue T
1	I-A	to White S/Blue non-T
2	I-B	to White S/non-Blue T
3	I-C	to White non-S/Blue T
4	I-D	to non-White S/Blue T
5	III-B-1	Additional response (within window)
6	III-C-1	Multiple response (within window)
7	III-A-1	Response to White S
8	III-A-2	Response to White non-S
9	III-A-3	Response to non-White S
10	LATE	Late response to White S/Blue T
11	II-A	to White S/Blue non-T
12	II-B	to White S/non-Blue T
13	II-C	to White non-S/Blue T
14	II-D	to non-White S/Blue T
15	III-B-2	Additional response (outside window)
16	III-C-2	Multiple response (outside window)

* These codes correspond to the codes which will appear in the **Reaction Time** printout.

TABLE 10 (A)

Block 1

	Orange U
	Yellow I
TGT	White S, Blue T
	Yellow L
	Green Y
B	White S, White T
	Purple F
	Brown A
	Magenta I
TGT	White S, Blue T
	Purple L
C	White P, Blue T
	Yellow U
	Brown U
B	White S, Yellow T
	Purple U
	Magenta I
D	Magenta S, Blue T
	Magenta L
B	White S, Orange T
	Orange C
TGT	White S, Blue T
	Yellow I
	Green F
A	White S, Blue P
	Orange H
D	Green S, Blue T
	Purple P
C	White U, Blue T
	Brown C
A	White S, Blue L
	Orange H
	Orange E
TGT	White S, Blue T
	Purple U
C	White C, Blue T
	Green O
A	White S, Blue L
	Magenta U
B	White S, Brown T
	Purple F
TGT	White S, Blue T
	Green P
C	White P, Blue T
	Magenta U
	Green F
TGT	White S, Blue T
	Yellow L
D	Green S, Blue T
	Magenta I
TGT	White S, Blue T
	Brown U
B	White S, White T
	Magenta I
TGT	White S, Blue T
	Magenta U
A	White S, Blue Y
	Brown Y
TGT	White S, Blue T
	Magenta L
D	Blue S, Blue T
	Magenta E
C	White A, Blue T
	Green I
TGT	White S, Blue T
	Yellow A
A	White S, Blue U
	Orange U
D	Purple S, Blue T
	Purple H

Block 2

	Yellow U
	Magenta I
C	White Y, Blue T
	Magenta E
TGT	White S, Blue T
	Purple O
B	White S, Orange T
	Brown F
	Green L
C	White E, Blue T
	Orange A
TGT	White S, Blue T
	Green A
A	White S, Blue P
	Brown C
	Green H
TGT	White S, Blue T
	Magenta Y
D	Brown S, Blue T
	Purple F
TGT	White S, Blue T
	Brown Y
D	Yellow S, Blue T
	Orange P
	Magenta L
C	White I, Blue T
	Green Y
TGT	White S, Blue T
	Magenta A
D	Blue S, Blue T
	Yellow U
C	White C, Blue T
	Orange O
B	White S, Purple T
	Brown E
C	White A, Blue T
	Yellow Y
A	White S, Blue H
	Orange U
	Green O
B	White S, Purple T
	Orange H
TGT	White S, Blue T
	Purple U
A	White S, Blue I
	Magenta H
B	White S, White T
	Magenta I
	Brown C
TGT	White S, Blue T
	Purple F
A	White S, Blue P
	Brown H
TGT	White S, Blue T
	Brown P
D	Brown S, Blue T
	Purple P
B	White S, Yellow T
	Green P
	Purple U
D	Yellow S, Blue T
	Magenta H
TGT	White S, Blue T
	Green Y
A	White S, Blue O
	Purple C
TGT	White S, Blue T
	Magenta F
	Yellow U
	Yellow E

TABLE 10 (B)

Block 3

TGT White S, Blue T
Yellow U

A White S, Blue H
Orange C

TGT White S, Blue T
Green E
Magenta I

D Purple S, Blue T
Yellow O

TGT White S, Blue T
Brown U
Yell[illegible] I
Green H

C White C, Blue T
Brown I

TGT White S, Blue T
Yellow H

B White S, Magenta T
Orange H

D Magenta S, Blue T
Magenta F

B White S, Orange T
Purple I

TGT White S, Blue T
Orange E
Green L

D Brown S, Blue T
Purple P

A White S, Blue P
Pur[illegible] H

D Magenta S, Blue T
Brown Y

A White S, Blue P
Orange U

TGT White S, Blue T
Purple C

C White L, Blue T
Yellow I
Magenta Y
Orange E

A White S, Blue E
Orange F

C White H, Blue T
Brown P
Yellow I

TGT White S, Blue T
Green H

B White S, Purple T
Orange A

A White S, Blue U
Brown U

B White S, White T
Orange I

TGT White S, Blue T
Green E

C White H, Blue T
Magenta U

TGT White S, Blue T
Purple L
Green O

B White S, Purple T
Orange C
Orange I

TGT White S, Blue T
Yellow H
Orange Y

C White O, Blue T
Green A

D Orange S, Blue T
Purple A

Block 4

Brown [illegible]

C White [illegible], Blue T
Magenta [illegible]

TGT White S, Blue T
Orange [illegible]

C White Y, Blue T
Yellow [illegible]

A White S, Blue Y
Purple [illegible]

TGT White S, Blue T
Orange [illegible]

B White S, Brown T
Yellow F

TGT White S, Blue T
Orange P

C White Y, Blue T
Magenta P
Yellow P

D Brown S, Blue T
Green L

TGT White S, Blue T
Green I
Orange F

B White S, Brown T
Magenta C

D Orange S, Blue T
Brown C

C White Y, Blue T
Green U
Yellow Y

TGT White S, Blue T
Orange L

A White S, Blue S
Purple L

TGT White S, Blue T
Yellow C

C White P, Blue T
Brown I
Green C

TGT White S, Blue T
Brown O
Magenta E
Purple H

B White S, Magenta T
Yellow I

A White S, Blue S
Purple P

B White S, Orange T
Brown P
Yellow L

D Brown S, Blue T
Magenta H

B White S, Purple T
Magenta E

TGT White S, Blue T
Orange P
Brown P

D Green S, Blue T
Green Y

TGT White S, Blue T
Brown C

A White S, Blue C
Yellow E

TGT White S, Blue T
Purple I

D Orange S, Blue T
Magenta H

A White S, Blue F
Yellow L
Brown P

124

TABLE 10 (C)

Block 5

Magenta O
D Orange S, Blue T
Green F
Yellow H
TGT White S, Blue T
Magenta P
Orange Y
Green H
B White S, Magenta T
Magenta O
TGT White S, Blue T
Brown F
D Purple S, Blue T
Brown Y
A White S, Blue L
Brown A
TGT White S, Blue T
Green C
C White H, Blue T
Purple P
TGT White S, Blue T
Brown Y
C White Y, Blue T
Orange E
B White S, Magenta T
Magenta O
D Yellow S, Blue T
Green E
TGT White S, Blue T
Purple O
Brown A
A White S, Blue E
Yellow H
TGT White S, Blue T
Green L
D Orange S, Blue T
Brown P
B White S, Yellow T
Purple C
A White S, Blue A
Yellow H
D Orange S, Blue T
Purple C
A White S, Blue F
Green F
C White Y, Blue T
Purple U
TGT White S, Blue T
Brown E
A White S, Blue I
Green Y
Purple P
TGT White S, Blue T
Magenta C
B White S, Purple T
Orange A
Brown I
TGT White S, Blue T
Green A
C White E, Blue T
Magenta U
Yellow E
B White S, Brown T
Orange O
C White A, Blue T
Yellow I
Purple E
TGT White S, Blue T
Green Y
Orange I

Block 6

Purple I
TGT White S, Blue T
Brown A
A White S, Blue I
Yellow I
Purple H
D Brown S, Blue T
Purple P
TGT White S, Blue T
Orange E
Green H
C White U, Blue T
Magenta Y
A White S, Blue O
Magenta L
B White S, Purple T
Yellow A
A White S, Blue A
Orange F
D Purple S, Blue T
Brown U
Green F
C White Y, Blue T
Purple U
TGT White S, Blue T
Yellow H
C White I, Blue T
Green P
TGT White S, Blue T
Brown O
D Blue S, Blue T
Yellow P
TGT White S, Blue T
Orange A
B White S, Green T
Yellow C
Purple Y
TGT White S, Blue T
Brown I
A White S, Blue E
Magenta Y
Green L
TGT White S, Blue T
Yellow O
B White S, Yellow T
Orange U
TGT White S, Blue T
Brown C
Magenta F
C White F, Blue T
Yellow U
TGT White S, Blue T
Yellow H
A White S, Blue Y
Green F
Brown I
B White S, Purple T
Yellow A
Brown H
D Orange S, Blue T
Magenta E
TGT White S, Blue T
Brown O
C White T, Blue T
Purple E
D Green S, Blue T
Purple O
Purple P
B White S, White T
Orange I

125

TABLE 10 (D)

Block 7

C White H, Blue T
Yellow L
B White S, Green T
Magenta P
D Brown S, Blue T
Brown U
IGI White S, Blue T
Yellow U
Yellow H
B White S, Purple T
Orange P
A White S, Blue I
Green E
B White S, Orange T
Purple U
IGI White S, Blue T
Yellow U
Green I
B White S, White T
Orange L
IGI White S, Blue T
Brown U
C White H, Blue T
Magenta L
A White S, Blue E
Green F
D Green S, Blue T
Purple P
Orange H
A White S, Blue A
Yellow I
IGI White S, Blue T
Brown A
C Green S, Blue T
Yellow O
D White O, Blue T
Magenta U
Green Y
IGI White S, Blue T
Orange I
Purple U
C White P, Blue T
Yellow E
Magenta I
A White S, Blue H
Green C
IGI White S, Blue T
Purple A
A White S, Blue H
Magenta L
IGI White S, Blue T
Yellow F
Green I
C White I, Blue T
Brown A
IGI White S, Blue T
Green U
Purple F
Yellow H
D Magenta S, Blue T
Brown I
B White S, Yellow T
Magenta A
IGI White S, Blue T
Orange E
D Yellow S, Blue T
Purple L
IGI White S, Blue T
Green U
Yellow P

Block 8

Brown C
D Yellow S, Blue T
Green U
C White T, Blue T
Orange E
IGI White S, Blue T
Magenta H
Purple Y
B White S, Magenta T
Brown U
IGI White S, Blue T
Brown Y
A White S, Blue C
Green P
IGI White S, Blue T
Green F
A White S, Blue O
Yellow F
IGI White S, Blue T
Green U
D Purple S, Blue T
Yellow A
Yellow O
C White Y, Blue T
Magenta H
Purple E
B White S, Magenta T
Magenta F
C White L, Blue T
Purple F
B White S, Orange T
Orange A
IGI White S, Blue T
Purple Y
Brown I
C White A, Blue T
Purple U
A White S, Blue C
Magenta L
IGI White S, Blue T
Magenta E
Yellow A
C White C, Blue T
Purple P
IGI White S, Blue T
Orange O
D Brown S, Blue T
Green C
A White S, Blue S
Orange H
IGI White S, Blue T
Yellow A
D Blue S, Blue T
Magenta O
B White S, Green T
Green O
D Orange S, Blue T
Yellow H
IGI White S, Blue T
Green P
Magenta U
Purple C
B White S, Purple T
Yellow O
Orange F
A White S, Blue U
Yellow L
IGI White S, Blue T
Yellow H
Purple E

126

TABLE 10 (E)

Block 9

Green O
B White S, Orange T
Purple I
Yellow C
Magenta A
TGT White S, Blue T
Orange L
Yellow O
A White S, Blue O
Brown C
TGT White S, Blue T
Yellow H
D Magenta S, Blue T
Green P
A White S, Blue C
Yellow U
Brown O
D Purple S, Blue T
Green Y
B White S, Green T
Green E
A White S, Blue E
Purple Y
TGT White S, Blue T
Magenta U
B White S, Green T
Purple H
TGT White S, Blue T
Green C
C White F, Blue T
Magenta C
TGT White S, Blue T
Purple U
C White C, Blue T
Yellow U
B White S, Magenta T
Magenta L
C White L, Blue T
Orange P
A White S, Blue S
Magenta E
TGT White S, Blue T
Green E
Yellow C
C White I, Blue T
Orange P
Magenta A
D Orange S, Blue T
Orange O
TGT White S, Blue T
Purple I
A White S, Blue P
Orange E
C White L, Blue T
Purple F
Green P
D Green S, Blue T
Yellow O
TGT White S, Blue T
Purple O
B White S, Magenta T
Yellow A
TGT White S, Blue T
Brown H
Magenta Y
D Yellow S, Blue T
Orange F
TGT White S, Blue T
Purple O
Green I

Block A (10)

D Orange S, Blue T
Yellow H
C White E, Blue T
Green A
Magenta L
Brown U
TGT White S, Blue T
Yellow C
Orange O
Green C
B White S, White T
Yellow L
TGT White S, Blue T
Purple H
C White F, Blue T
Orange P
TGT White S, Blue T
Yellow U
Brown U
C White I, Blue T
Green O
D Yellow S, Blue T
Orange O
Brown F
C White S, Blue C
Orange F
TGT White S, Blue T
Orange H
D Blue S, Blue T
Orange U
B White S, Magenta T
Orange F
TGT White S, Blue T
Purple E
B White S, Orange T
Green Y
Magenta L
A White S, Blue O
Brown E
Green E
Yellow F
TGT White S, Blue T
Purple U
C White C, Blue T
Purple E
A White S, Blue I
Orange O
B White S, Orange T
Magenta O
D Yellow S, Blue T
Brown H
A White S, Blue L
Orange U
TGT White S, Blue T
Orange I
A White S, Blue E
Magenta F
B White S, Green T
Orange F
TGT White S, Blue T
Orange P
C White Y, Blue T
Green C
TGT White S, Blue T
Yellow L
D Yellow S, Blue T
Yellow Y
TGT White S, Blue T
Orange E
Brown A

127

COMPARING CLASSROOM AND CLINIC MEASURES OF ADD-H: DOSE RESPONSE EFFECTS

MARK D. RAPPORT, GEORGE J. DUPAUL, KEVIN KELLY, AND JAY JONES

THE CHILDREN'S LEARNING CLINIC
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF RHODE ISLAND
KINSTON, RI 02881

PAPER PRESENTED AT THE AMERICAN PSYCHOLOGICAL ASSOCIATION
LOS ANGELES, CALIFORNIA, AUGUST, 1985

Drug Studies

Abstract

Attention Deficit Disorder with Hyperactivity (ADD-H) is characterized by an inability to sustain attention, impulsivity, and overactivity (DSM-III; American Psychiatric Association, 1980). Related problems such as social and academic difficulties are frequently reported by parents and teachers (Aman, 1984). Although some follow-up studies have shown a reduction in untreated ADD-H children's symptomatology over the course of three to four years (Charles & Schain, 1981; Satterfield, Satterfield, & Cantwell, 1981), a corresponding improvement in academic functioning and social adjustment has not been demonstrated (Satterfield, Hoppe, & Schell, 1982).

The attentional component of the disorder is especially problematic owing to its interferring effects on school performance, home behavior and persistence into adult life, and has been empirically validated in both laboratory (Sykes, Douglas, & Morgenstern, 1973) and classroom settings (Rapport, Murphy, & Bailey, 1980; 1982). Thus far, the Continuous Performance Test (CPT; Rosvold, Mirosky, Sarason, Bransome, & Beck, 1956) has been the most frequently used laboratory test for identifying the attentional deficiencies in ADD-H children (Klee & Garfinkel, 1983). The CPT is an experimenter-paced vigilance task sensitive to momentary lapses of attention and yields two dependent measures: omission and commission errors; indicative of attentional and impulsivity problems, respectively (Sostek, Buchsbaum, & Rapoport, 1980). ADD-H children typically make fewer correct detections (i.e., omission errors) and more incorrect responses (i.e., commission errors) to presented stimuli compared to normal children.

Psychostimulant medication is presently the most widely used treatment for ADD-H (Safer & Krager, 1983), most likely due to its cost-efficiency and

demonstrated short-term effects on sustained attention (Sykes et al., 1973), activity level (Porrino et al., 1983), academic performance (Rapport, Stoner, DuPaul, Birmingham, & Tucker, 1985), and impulsivity (Brown & Sleator, 1979; Rapport, DuPaul, Stoner, Birmingham, & Masse, in press). Despite the findings of increased attention in ADD-H children in both school (Rapport et al., 1982; 1985) and laboratory (Sostek et al., 1980) environments following psychostimulant treatment, there is a dearth of information as to whether drug-related changes in these two settings are similar and/or dose specific.

The primary purpose of the present study was to investigate the relationship between the most frequently used laboratory method of assessing attentional deficits (CPT) and ADD-H children's attention in school under various doses of methylphenidate. A second purpose was to subject the expected dose-response curves to statistical analysis and determine whether a similar quadratic relationship (inverted U-shaped curve) is found between performance and dose as reported in previous studies (Brown & Sleator, 1979; Sprague & Sleator, 1977). Finally, a comparison of fixed-dose and milligram- per-kilogram (mg/kg) dose-response curves was planned to determine the relative advantage of each in assessing drug responsivity. Collectively, these results should have a direct bearing on assessing ADD-H children's responsiveness to psychostimulant medication and the putative ability of laboratory-based testing for establishing optimal dose.

METHOD

Subjects

Families of children with complaints of inattentiveness, impulsivity, and overactivity were referred to the Children's Learning Clinic (CLC) of the University of Rhode Island by pediatricians and local school system personnel. All children met each of the following diagnostic criteria: (a) an independent diagnosis by the child's pediatrician and CLC's directing clinical psychologist using DSM-III (American Psychiatric Association, 1980) criteria for Attention Deficit Disorder with Hyperactivity (ADD-H); (b) a maternal report of a developmental history consistent with ADD-H (Barkley, 1981); (c) a maternal rating of at least two standard deviations above the mean for the child's age on the Werry-Weiss-Peters Activity Scale (WWPAS: Routh, Schroeder, & O'Tuama, 1974) and problems in at least 50% of the situations on Barkley's (1981) Home Situations Questionaire; (d) a teacher rating on the Abbreviated Conners Teacher Rating Scale (ACTRS) above 15, the designated cutoff score for hyperactivity (Werry, Sprague, & Cohen, 1975); (e) performance on the Matching Familiar Figures Test (MFFT: Kagan, Rosman, Day, Albert, & Phillips, 1964) characteristic of a "fast-inaccurate or impulsive" responder (i.e., faster than average responses and higher than average error rates for the child's age); and (f) absense of any gross neurological, sensory, or motor impairment as determined by pediatric examination.

A total of 24 children (20 males, 4 females) between the ages of 7 and 10 years (M = 8.8 years) participated in the study after informed consent was obtained from their parents. All children were of average or above average intelligence (mean IQ = 98.3, SD = 6.6) as assessed by the Peabody Picture Vocabulary Test-Revised, Form L (Dunn & Dunn, 1981) and from families of low to middle socioeconomic status (Hollingshead, 1975). No children were taking medication prior to participating in the study.

Experimental Design and Drug Administration

A double-blind, placebo control, within-subject (crossover) experimental design was used in which all children received each of the five doses (following baseline measures) in a randomly assigned, counterbalanced sequence.

Methylphenidate (MPH) was prescribed by each child's pediatrician in the following doses: placebo, 5 mg, 10 mg, 15 mg, and 20 mg. Fixed doses (vs mg/kg) were prescribed to reflect typical pediatric practice in the United States (Physicians' Desk Reference, 1985) and because children's response to MPH dosage manipulations has not been shown to be dependent upon total body weight (Kinsbourne & Swanson, 1979, p. 14; Rapport et al., 1985). All MPH and placebo dosages were packaged in colored gelatin capsules by the clinic's pharmacist to avoid detection of dose and taste. Capsules were placed in individual daily dated envelopes to insure accurate dose administration.

Children were seen once per week at the CLC for individual testing. Baseline measures were obtained during the child's second clinic visit, to allow familiarization with clinic personnel and testing procedures. On subsequent testing days, children were administered a capsule of either the active agent (MPH) or a placebo. Testing (under double-blind conditions) was begun 90 min after oral ingestion to insure optimal medication effect (cf. Kinsbourne & Swanson, 1979, p. 11).

Children were classified as favorable responders or non-responders to MPH in a manner similar to previous studies (Thurston, Sobol, Swanson, & Kinsbourne, 1979), using the Paired Associates Learning (PAL) test. The PAL test was administered once per week under each experimental condition described above. Those children showing a drug-induced facilitation in performance of 25% of more (i.e., a 25% drop in error rate) compared to baseline or placebo were classified as favorable responders, while those whose performance neither improved to the above criteria nor declined in relation to baseline or placebo were classified as non-responders. All children in the present study were found to be favorable responders using the above criteria.

Assessment

Clinic Measures

During each child's weekly session the Gordon Diagnostic System (GDS; Gordon & McClure, 1983) version of the Continuous Performance Test (CPT) was administered. In using this version, the child is presented with a series of numbers ranging from one through nine at 1 sec intervals and is required to press a button on the apparatus each time the number "1" is immediately followed by the number "9". Total correct responses, omission (failing to identify the "1-9" combination when it appears) and commission errors (pressing the button for digit combinations other than the correct combination) are recorded for the 9 min test duration.

Classroom Measures

The classroom phase of the experiment involved 6 consecutive weeks of assessment and observation (see Teacher Ratings and Behavioral Observations sections below) which corresponded with the 6 week clinic testing period decribed above. Following baseline data collection (1st week), parents were given a week's worth of medication in pre-dated envelopes at one dose level (i.e., placebo, 5 mg, 10 mg, 15 mg, or 20 mg MPH). This prodecure continued until a child received each dose for 6 consecutive days. All weekly dosage changes occurred on Sundays (no medication was administered on Saturday to allow for "washout") due to the inter-individual variation in serum and blood plasma levels following an acute administration of MPH (Gualtieri et al., 1982) and to control for potential rebound effects. Parents were instructed to give their child a medication capsule each morning, 1/2 hr prior to breakfast, and to tell them it was a vitamin supplement (i.e., to help control for expectancy effects, cf., Rapport et al., 1982). Both used and unused envelopes were returned to the CLC on a weekly basis to control for medication compliance. Weekly teacher ratings and daily classroom observations (described below) were completed during the morning hours due to the behavioral time-response course of MPH (Swanson, Kinsbourne, Roberts, & Zucker, 1978).

Teacher Ratings. Classroom teachers completed the ADD-H: Comprehensive Teacher Rating Scale (ACTeRS; Ullmann, Sleator, & Sprague, 1985) each Friday throughout the experimental conditions, which reflected the children's behavior during the morning hours (until 11:30 a.m.) of that week. Morning ratings were obtained due to the relatively short half-life of MPH (i.e., 2-3 hr). All teachers were blind as to when medication was administered and specific doses. The ACTeRS includes 24 items describing classroom behavior on four factors (attention, hyperactivity, social skills, and oppositional behavior), has adequate psychometric properties, and has been shown to be highly sensitive in detecting medication effects (Ullmann et al., 1985).

Behavioral Observations. Children were observed in their regular classrooms for 20 min intervals, 3 days per week across the 6-week evaluation period. Classroom observations began approximately 1 1/2 to 2 hr after the children's morning medication, during which time the class completed in-seat academic work assigned by the teacher. Children were unobtrusively observed by trained undergraduate and graduate students for 60 consecutive intervals during each observation period throughout the study. Each interval was divided into 15 s of observation followed by 5 s for recording. All observers were blind as to when medication was administered and specific doses.

A child's behavior was categorized as either on-task or off-task in a manner identical to that used by Rapport et al. (1985). Off-task behavior was defined as visual nonattention to one's materials for more than 2 consecutive s within each

15 s observation interval, unless the child was engaged in some other task-appropriate behavior (e.g., sharpening a pencil, asking the teacher a question)

Academic Measures. Children's performance on regularly assigned academic work during the scheduled observation periods was used as a dependent measure in an effort to preserve ecological validity, yet still maintain adequate experimental control. Academic seat-work typically involved completing arithmetic problems or language arts assignments. Assignments were graded after class by either the teacher or primary observer. Daily performance was recorded for both the percent of problems completed and the percent of problems completed correctly.

Reliability. An inter-observer reliability check of each child's on-task behavior was taken on 33% of the observation days, and at least once during each experimental phase. Overall reliability was consistently above 80%, with a mean of 88% across children. A mean Kappa value of .85 was obtained for all children.

Reliability checks of academic measures were completed daily for all children. Checks were made by either the primary observer in the classroom or the teacher (whomever did not initially score the child's paper for that day), and consisted of scoring the child's assignment independently of the other. Agreement was defined as agreement on the number of problems completed and number of problems completed correctly. Reliability was computed by dividing agreement between observer and teacher by agreement plus disagreement and multiplying by 100 to calculate the percentage. Observer-teacher agreement on problems completed and performance accuracy was consistently above 96% for all children across experimental conditions.

Results

A one-way multivariate analysis of variance (MANOVA) with repeated measures across dose (MPH) was performed using each of the six dependent variables (CPT omission and commission errors, ACTeRS attention factor scores, percent on-task, assignments complete, and assignments correct) and was highly significant ($F(30, 442) = 5.05$, $p < .001$) according to Wilk's lambda criterion. Follow-up univariate analyses of variance (ANOVA) with repeated measures across dose were performed separately for each dependent measure. Due to the number of planned follow-up tests, a Bonferoni critical value of .008 was established a priori to control for potential experimentwise error (i.e., univariate levels must exceed this value to be considered statistically significant).

Significant overall effects were found for CPT commission errors ($F(5,115) = 5.29$, $p < .001$), on-task behavior ($F(5, 115) = 23.64$, $p < .001$), percent of assignments completed ($F(5, 115) = 9.44$, $p < .001$), and ACTeRS attention factor scores ($F(5, 115) = 9.65$, $p < .001$). Percent of assignments completed correctly did not exceed Bonferoni critical values ($F(5, 115) = 3.19$, $p < .05$) and non-significant overall effects were found for CPT omission errors.

Newman-Keuls post hoc analyses were conducted for each of the significant univariate findings to examine between-dose differences on the various dependent measures. In the first of these analyses, the mean number of CPT commission errors (indicative of impulsive responding) under placebo was found to be significantly greater than those made under the 10 mg ($p < .05$), 15 mg and 20 mg ($p < .01$) conditions. A significantly greater number of commission errors were also made under 5 mg compared to the higher 15 mg ($p < .05$) and 20 mg ($p < .01$) doses. Similar between-dose differences were found for the length of time children spent attending to their academic assignments in the classroom. Children's on-task behavior was not significantly different during baseline and placebo conditions, but was significantly enhanced under the 10 mg, 15 mg, and 20 mg weekly MPH conditions ($p < .01$). Significantly higher on-task behavior was also found under the 10 mg ($p < .05$), 15 mg and 20 mg ($p < .01$) MPH conditions compared to 5 mg.

The mean number of weekly academic assignments completed by the children was significantly enhanced during the 10 mg, 15 mg, and 20 mg MPH conditions compared to baseline and placebo ($p < .01$), with the 10 mg ($p < .05$), 15 mg and 20 mg ($p < .01$) MPH doses also resulting in higher academic completion rates compared to the lower 5 mg dose.

A number of significant between-dose differences were found for the teacher ratings of children's classroom attention (ACTeRS attention factor). For example, children were rated as significantly more attentive under the 10 mg ($p < .05$), 15 mg and 20 mg ($p < .01$) weekly MPH conditions compared to baseline and placebo. They also received significantly higher ratings (indicative of improvement) under the 15 mg ($p < .05$) and 20 mg ($p < .01$) conditions compared to 5 mg, and under 15 mg and 20 mg ($p < .05$) compared to 10 mg. No further between-dose effects were found for any of the dependent measures.

In a subsequent analysis, tests for trend (Winer, 1969) were used to examine the shape of the relationship between dose (MPH) and each of the dependent measures. Although dependent variables showing non-significant univariate effects are not traditionally subjected to follow-up trend analysis, they were included herein (i.e., percent of assignments completed correctly and CPT omission errors) to determine whether a quadratic dose-response relationship (i.e., an inverted U-shaped function) may have precluded finding significant overall effects. A significant linear relationship between MPH dose and each of the dependent measures (percents on-task, assignments complete and correct, ACTeRS attention factor ratings, and CPT commission errors) was found except for CPT omission errors (see F values in Table 1), with all other potential relationships (e.g., quadratic, cubic) non-significant. Thus, the majority of dependent measures improved as a function of increasing dose in a linear fashion. A follow-up test for equality of slopes, with dose as the covariate and each dependent measure treated as a separate group, was non-significant (F (4, 590) = 1.31, N.S.), indicating that the significant linear relationships described above were of approximate equal slope and intercept.

The dose-response relationship between MPH and each of the six dependent measures was planned using both fixed-dose and milligram per kilogram (mg/kg) plotting methods. A comparison between the two methods was performed to address the question regarding whether response to MPH is dependent upon total body weight. The subsample of 6 children was randomly selected from the total N to facilitate graphical illustration and interpretation of individual dose responsiveness (i.e., drawn curves using all 24 children would be too difficult to visually interpret).

A series of t-tests was conducted to compare the dose-response relationship obtained for the subsample children (n = 6) with the total sample (N = 24) on each of the six dependent measures. No significant differences were found between the mean scores of the subsample and those of the total sample for each of these variables across dosages, indicating that the former were representative of the total group's response to MPH.

In examining the curves depicted in Figure 2 (Note: this Figure is presently being prepared), it appears that children's performance across the various dependent measures show similar dose-related changes using both fixed-dose and mg/kg plotting methods; thus lending support to the notion of individual responsivity to psychostimulant medication. Moreover, total body weight alone does not appear to adequately explain these differences, as several children with similar or identical weights (see mg/kg graphs in Figure 2) showed dramatically different response patterns under active drug conditions.

Although the response of most children to differing doses of MPH was consistent across tasks and behaviors in the present study, a few of the children evinced task-specificity. For example, two of the children showed continued gains or stability in percents on-task and assignments completed as a function of increasing dosage, however, both children's MFF test commission errors scores decreased at the higher dose levels.

Discussion

The results of the present investigation demonstrate a functional relationship between MPH and ADD-H children's attention in the classroom environment. As a group, children showed significant decreases in their CPT commission errors (indicative of reduced impulsivity) and corresponding increases in classroom attention (on-task behavior) and academic productivity across most MPH dosages compared to placebo. The dose-response relationship between the various dependent measures and increasing dosages of MPH was linear.

Curiously, the CPT's primary measure of attention, omission errors, was not significantly affected by any of the MPH dosages. Several explanations may have accounted for this result. For example, past investigators finding no significant drug effect on CPT omission errors reanalysed their data using only those children showing abnormal levels of omission errors during baseline conditions and found significant drug effects (Charles, Schain et al., 1979). To investigate this possibility, we conducted a similar post-hoc analysis using the 20 children showing abnormal baseline omission error scores (based on GDS standardization criteria), yet continued to find no drug effect. Given the significant dose-response effect of MPH on the other dependent measures, a more likely explanation of the results is that the GDS version of the CPT is not a medication sensitive measure for detecting changes in ADD-H children's attention (Swanson, Barlow, & Kinsbourne, 1979). Previous investigations have found similar results using both ADD-H children (Campbell, Douglas, & Morgenstern, 1971; Rapoport et al., 1980) and normal adults (Klorman, Bauer et al., 1984) treated with psychostimulants

The continued improvement in children's classroom attention and academic performance at higher dose levels was surprising, given that previous investigations have generally shown a deterioration of performance at doses exceeding 0.3 mg/kg (Brown & Sleator, 1979; Sprague & Sleator, 1977). At least two factors may have been responsible for these differences. The performance decrements at higher dosages reported in previous studies have largely occurred on laboratory-based tests of short-term memory (Sprague & Sleator, 1977) and impulsivity (Brown & Sleator, 1979), which may be more sensitive to drug effects This explanation appears unlikely, however, as recent studies have shown increases in ADD-H children's performance on similar tests of short-term memory (Rapport et al., 1985), impulsivity (Rapport, Stoner, DuPaul et al., 1985), and attention (Rapport, DuPaul, Stoner, et al., in press) with dosages exceeding 0.3 mg/kg. A more likely explanation involves Sprague and Sleator's (1975) earlier hypothesis of continued cognitive improvement beyond 0.3 mg/kg, which could not be tested by the interpolated dose-response curve drawn between placebo, 0.3 mg/kg, and 1.0 mg/kg in the 1977 and 1979 studies. Thus, children may continue to evince improvement beyond 0.3 mg/kg, with performance decrements occurring as they approach 1.0 mg/kg.

In addition to the overall findings discussed above, a molecular analysis of individual children's response to MPH revealed several interesting results. In comparing the fixed-dose to mg/kg response curves for example, it was apparent that both methods of data plotting yield similar functional relationships between behavior and dose. Of particular interest was the idiosyncratic response children exhibited for a particular dependent measure (e.g., percent of academic assignments completed) despite identical or similar body weights.

Additional within-subject differences were apparent for a particular dose across behavioral domains. Although a clear majority of children showed consistent changes across the various dependent measures as a function of dose, the response of others to differing doses of MPH was task-specific (i.e., optimal responsivity depended not only on a child's reaction to a particular dose, but also on the behavior being assessed).

AUTHOR'S NOTE

The results presented herein are part of a larger clinical dose-response study presently being conducted by the Children's Learning Clinic to examine ADD-H children's CPT and academic performance. There will be approximately 36 children involved at the study's conclusion. Consequently, the results presented above may change, as would any interpretations offered to explain the results. We respectively request that no person cite the findings presented above without specific written permission of the authors. A complete manuscript of the study's results will be prepared upon completion of data collection. Appropriate graphs and figures will also be available at that time.

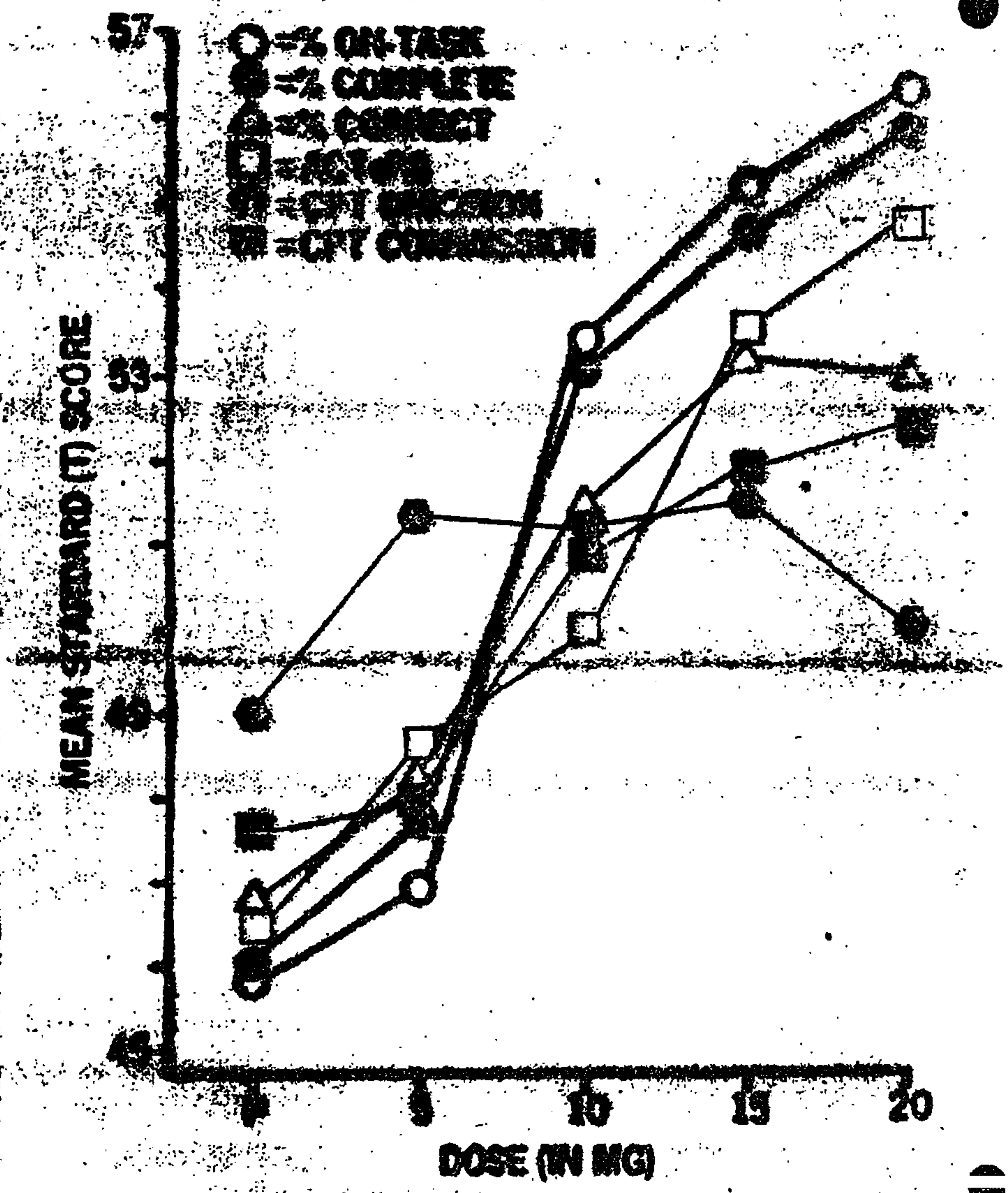
TESTS for TREND (with baseline omitted from analysis)

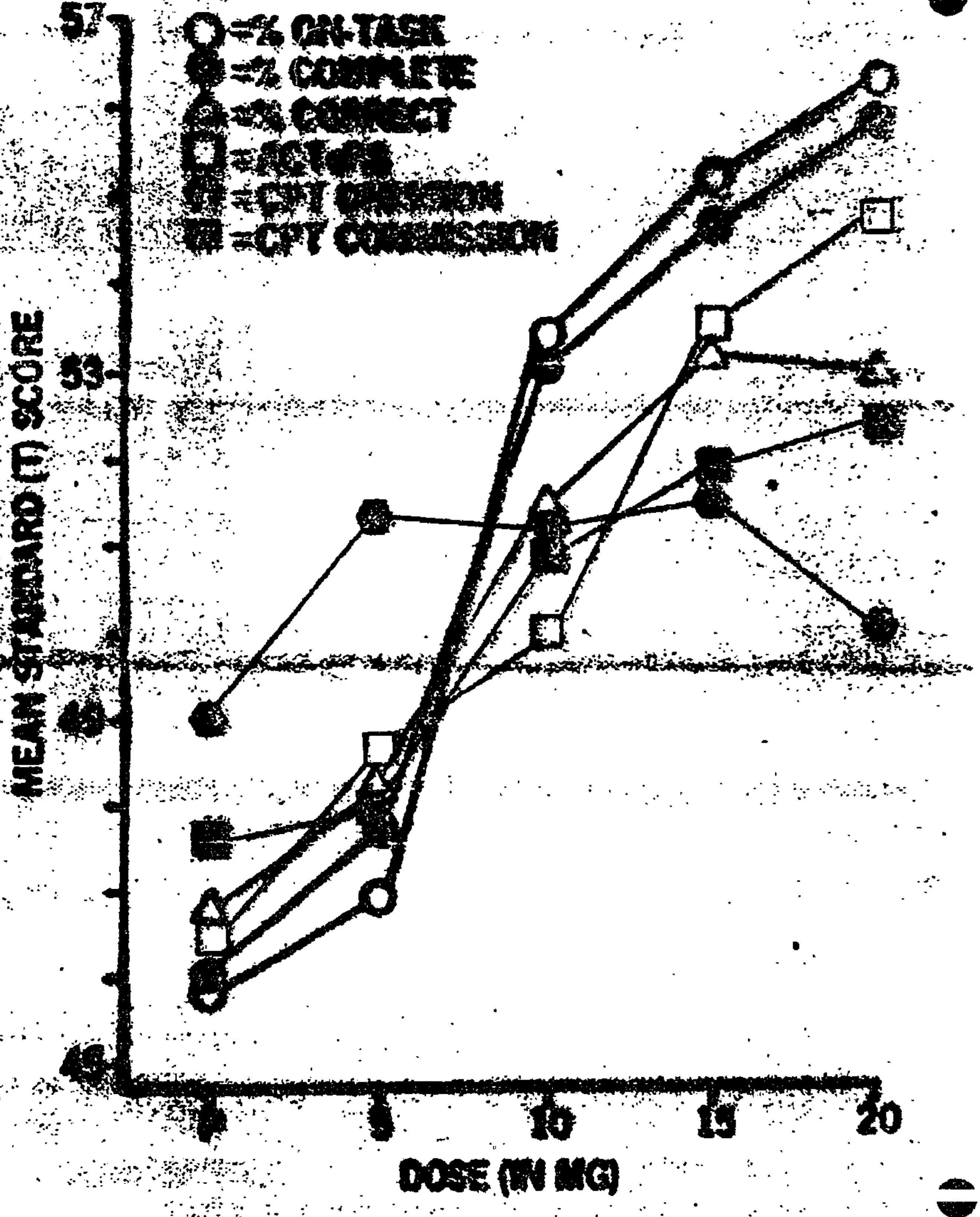
F Values for Trend Analysis

Measure	Linear	Quadratic	Cubic	Quartic	df
	FUNCTION				
% On-task	62.22***	1.14	2.74	2.06	1,92
% Complete	33.24***	.76	.74	.75	1,92
% Correct	13.79***	.81	.84	.09	1,92
ACTeRs	31.95***	.04	.24	.54	1,92
Omission Errors	46	4.06	.06	.33	1,92
Commission Errors	29.43***	.39	1.46	.80	1,92

***$p < .001$

Test for Equality of Slopes (ANCOVA with dose as the covariate and each measure treated as a separate group):
$F(4,590) = 1.31$, N.S.


=% ON-TASK
=% COMPLETE
57
53
49
45
MEAN STANDARD (T) SCORE
0
5
10
15
20
DOSE (IN MG)


=% ON-TASK
=% COMPLETE
=% CORRECT
=ACTeRS
=CPT OMISSION
=CPT COMMISSION
MEAN STANDARD (T) SCORE
57
53
49
45
DOSE (IN MG)
5
10
15
20

144

Development of a Multi-Method Clinical Protocol for Assessing Stimulant Drug Response in ADD Children

Russell A. Barkley

Department of Psychiatry

University of Massachusetts Medical Center

and

Mariellen Fischer, Robert F. Newby, and Michael J. Breen

Department of Neurology

Medical College of Wisconsin

Running Head: Stimulants and ADD

Address for Reprints: Russell A. Barkley, Ph.D., Department of Psychiatry, University of Massachusetts Medical Center, 55 Lake Avenue North, Worcester, MA 01605.

Abstract

A protocol was developed for the clinical assessment of stimulant drug response in children with Attention Deficit Disorders (ADD) that is easy to administer. A triple-blind, placebo-controlled crossover design was used to assess two dose of Ritalin (0.3 mg/kg and 0.5 mg/kg BID) and a placebo. Each condition lasted one week. A one hour assessment battery was administered at the end of each drug condition. This consisted of parent and teacher ratings of ADD behaviors, problem settings, drug side-effects, laboratory tests of vigilance and impulse control, and clinic playroom observations of ADD behaviors during academic performance. Twenty-three of 28 children (ages 5 to 12 years) completed the entire study. Results demonstrated that most measures in the protocol were sensitive to both doses of medication. Eighty percent of the children were judged positive responders to medication. While minor revisions to the present protocol are suggested, it is clear that a quantitative, objective, cost-effective protocol for the clinical evaluation of stimulant drug response is feasible and should represent a substantial improvement over the subjective impressions now commonplace in clinical practice. Local centers equipped with the necessary resources for performing this type of objective evaluation should assist physicians by providing a standardized protocol for the early identification and rational pharmacological treatment of this significant mental health problem.

146

Development of A Multi-Method Clinical Protocol for Assessing Stimulant Drug Response in ADD Children

Stimulant drugs remain the most commonly recommended treatment for children with Attention Deficit Disorders (ADD), particularly in pediatric and child psychiatry clinics (Barkley, 1981, Dulcan, 1984). Yet, the decision to medicate such children is often based solely upon initial parental interviews and the titration of drugs solely on telephone conversations during the drug trial. While many decry the unreliability of such sources of information in general (Ullman, Egan, Fiedler, Jurenac, Pliske, Thompson, & Doherty, 1981) and point to the need for more systematic and objective evaluations of drug responding specifically (Cantwell, 1980; Ross & Ross, 1982), an objective, multi-method, yet cost-effective assessment protocol for use in clinical practice has remained elusive. Objective laboratory tests and direct observational methods for evaluating stimulant drug effects have been used for more than 15 years in the research literature (Barkley, 1977; Cantwell & Carlson, 1978) but are expensive, cumbersome, or unavailable to clinicians. However, recent developments in direct behavioral observation systems (Barkley, 1983a; Milich, Loney, & Landau, 1982) and micro-computer technology (Gordon, 1983; Klee & Garfinkel, 1983) allow for the adaptation of these previously cumbersome research instruments to the efficient clinical assessment of inattention, impulsivity, and other ADD behaviors in children.

This paper describes the development of a protocol for the clinical assessment of ADD children and their stimulant drug responses and its use with 28 clinic-referred ADD children. Varley and Trupin (1983) previously described a double-blind clinical protocol for evaluating stimulant drug responding but this procedure employed only brief parent and teacher ratings of behavior. Kinsbourne and Swanson (See Dalby, Kinsbourne, Swanson, &

Sobol, 1977) were probably the first to report the use of an objective instrument for clinically assessing stimulant drug responses, but they relied solely on one measure (paired-associate learning task) of only one type of behavior and typically conducted the evaluation of the child within a single day. More reliable and valid information can be obtained from assessments that rely on different sources of information (parent, teacher, clinician) from different settings (home, school, community), using different methods (ratings, laboratory tests, direct behavioral observations) combined into a single protocol (Mash & Terdal, 1981) which examines drug responding over a reasonable length of time. After reviewing the recent research literature, the present authors selected those methods demonstrating the most reliable sensitivity to stimulant drug effects while offering a minimum investment of time and resources for a clinical service.

The Drug Evaluation Protocol

Procedures

After receiving approval from the Human Research Review Committee of the Medical College of Wisconsin, the assessment service was advertised by letters sent to all pediatricians, child psychiatrists, and child neurologists in the greater Milwaukee area. Care was taken to explain the purpose and procedures of the service and to assure area clinicians that ongoing professional care of the ADD children would remain under their supervision. Clinicians wishing to refer children to the service provided a letter of referral, a separate letter authorizing school staff to administer the noon dose of medication (mandatory in Wisconsin public schools), and a prescription for both doses of methylphenidate (0.3 and 0.5 mg/kg given twice daily).

This prescription was sent to the pharmacy of an affiliated hospital for preparation of the two doses and placebo. The lactose placebo and

148

appropriate dosages were prepared to the nearest 2.5 mg using regular Ritalin tablets (5 and 10 mg size) which were then crushed and placed within orange opaque gelatin capsules (size #0, Ely Lilly Co.). Capsules disguised both the differences in doses across the three drug conditions as well as the distinct taste differences between Ritalin and placebo. While each drug condition lasted 1 week, extra capsules were provided to permit the rescheduling of the clinic visits when necessary (e.g., child illness) without undue inconvenience in supplying families with more medication. Unused capsules were returned to the clinic staff at the end of each week to allow for a convenient check on compliance to the drug regimen.

Children were initially seen for a 1.5 hour evaluation consisting of the measures described below. This screening was conducted to insure that children referred for this service were appropriate for such a drug trial. Of 28 children between 5 and 12 years of age referred to the service during its first 9 months, 3 were deemed inappropriate for the drug trial based upon this initial evaluation. One child had a history of multiple tic disorder and was rejected because of possible increased tics with stimulants (Comings & Comings, 1984; Golden, 1982); a second child had a history of a cerebro-vascular accident 1 year earlier and was rejected because of the known cardiovascular pressor effects associated with Ritalin (Cantwell & Carlson; Hastings & Barkley, 1978); and a third child was diagnosed as having only Oppositional Disorder of Childhood without evidence of ADD.

Those children deemed eligible for the protocol were then scheduled for 3 weekly evaluations and their parents provided with the first week's supply of the drug/placebo as well as the telephone numbers of the first author. If parents observed side effects of concern to them, they were instructed to contact the first author who would then notify the referring physician and discuss whether or not to discontinue the drug trial. Two of the 28 children initially referred were discontinued during the drug evaluation due to the development of tic reactions to the medication. Children were assigned to one of six possible drug orders of the placebo (P), low dose (L), and high dose (H) conditions: PLH, PHL, LPH, LHP, HLP, HPL.

The parents, children, their teachers, and the clinical assistant conducting the weekly drug evaluations were kept blind to the child's drug condition and order until the end of the entire evaluation. At that time, the drug code was broken and parents informed of the order and dose levels. The children were always evaluated on the last day of each drug condition and the clinical assistant insured that the children received a dose of Ritalin within 1 hour of the evaluation. After the final assessment, the findings of the entire evaluation were then reported to the referring physician and care of the child returned to his/her clinical practice. Recommendations for additional forms of treatment (e.g., parent training in behavior management, marriage counseling, individual therapy for parents evidencing depression, referral of child for special educational assessments, etc.) were occasionally made based upon the findings of the initial clinic evaluation of the children and their parents. Many of these therapies were also available at the clinic where the initial evaluations were conducted.

Assessment Methods

Initial Clinical Evaluation. The initial clinical evaluation consisted of the measures described below:

1. Semi-structured parental interview regarding: present behavioral, emotional, and learning problems; developmental, medical, and school histories; and family history and current family problems.

2. Vigilance and Impulse Control: These constructs were assessed using the Gordon Diagnostic System (GDS: Gordon, 1983). This system used a small computer that permitted two tests to be administered to the child in a machine-paced procedure. The device was a metal box with a display screen on the front surface and a large blue button beneath it. Just above the screen were one red and one green light. The first test was a Delay Task believed to assess impulse control. It required that the child sit before the device, wait a short period of time, and then press the blue button. If the child waited a sufficient period of time (6 seconds) and then pressed the button, the green light appeared and the child earned a point. Cumulative points were displayed on the screen. If the child did not wait an adequate period of time before pressing the blue button, the green light did not appear and no point was awarded. The children were not told bu the experimenter how long to wait -- only that if they waited long enough they could earn a point. the task lasted 8 minutes and the child's scores were the number of correct responses (successful delays), the total number of responses, and the ratio of correct to total responses (Efficiency Ratio). Gordon (1979) has shown that the Delay Task significantly discriminates hyperactive from non-hyperactive clinic-referred children and is significantly correlated with teacher ratings from the Conners Teacher Rating Scale.

The second test was a measure of vigilance or sustained attention, similar to the continuous performance tasks used in stimulant drug research with ADD children (Barkley, 1977). The child sat before the device and a series of numbers were shown on the screen at a rate of one per second. The 6 to 12 year old children were to press the blue button whenever the 1-9 number sequence appeared. The 5 yaer old children were told to watch only for the number 1 to appear and then to depress the button. The task lasted 9

151

minutes. The child's scores were the number of correct responses, number of omissions (missed target stimuli), and number of commission errors. Normative data are available for children ages 3 to 16 years for both GDS tasks.

3. ADD Behaviors During A Restricted Academic Situation: Following the testing of the children, the experimenter took the mother and child to a playroom (12 x 18 ft.) equipped with a one-way mirror and intercom. The mother was asked to have her child perform a series of math problems. The child sat at a small table away from his/her mother. The mother was seated on a sofa situated across the room and read magazines. The mother was told to see to it that the child remained at the table, did not interrupt her, and completed the problems. Three sets of math problems with varying difficulty levels were available. The problems were chosen from commonly available math workbooks at local educational stores. That set most appropriate for the child's age was used. Sets were available for ages 5-6, 7-8, and 9-11 years. Each set contained 5 pages of math problems to insure that the child did not finish the task before the 15 minutes of observation were completed. Three different versions of each difficulty level were used to reduce the likelihood of practice effects over the 3 weeks.

During this 15 minute period, the clinical assistant coded mother and child behavior. A tape recorder was used to cue the coder to the occurrence of every 30 second interval. During each interval, the coder scored the occurrence of each of eight behaviors. The category was checked as occurring only once regardless of the frequency of its occurrence during that 30 second interval. The categories were: Off-Task, Fidgets, Vocalizes, Talks to Mother, Plays with Objects, Out-of-Seat, Child Negative, and Mother Commands. A single score was derived for each category, this being the percentage occurrence of these behaviors relative to the total possible occurrences for that behavior. A total percent occurrence for all behaviors was also derived as a single Total ADD Behaviors score.

This observational system was selected to evaluate those disruptive behaviors most frequently seen in ADD children during situations many ADD children find most difficult; situations in which they are required to accomplish sustained tasks independently of adult assistance. Prior research has shown that a similar restricted academic situation and coding system not only discriminated ADD from normal children but also from non-ADD conduct problem children (Milich, Loney, & Landau, 1982). A similar coding system also showed significant stability over a 2 year period for normal and clinic-referred children (Milich, 1984; Milich, Loney, Whitten, 1983), and significant correlations with teacher ratings of hyperactivity and behavior problems.

The coder was trained to a level of 75 percent agreement with the first author using pre-recorded videotapes of mothers and children in this situation. Subsequently, intercoder reliability estimates were taken on five of the subjects by a second coder. Intercoder reliability was calculated using the number of agreements divided by the total possible number of occurrences. Inter-coder reliability averaged .77.

4. Parent Ratings of Child Behavior: While the child was being given the laboratory tasks, the mother completed several rating scales of her child's behavior. These were:

a. Child Behavior Checklist (CBCL; Achenbach & Edelbrock, 1983): This is a 118 item rating scale having two sections, one assessing Social Competence and the other Behavior Problems. The scale was used because it provides a generally comprehensive assessment of parent ratings of most of the commonly occurring dimensions of psychopathology in children and has been shown to discriminate ADD from normal children (Edelbrock, Costello, & Kessler, 1984). Normative data are available for ages 4 to 16 years for both sexes.

b. Conners Parent Rating Scale - Revised (CPRS-R; Goyette, Conners, & Ulrich, 1978): This is a 48 item questionnaire concerning various types of child behavior problems. The scale can be scored to yield five factors: Conduct Problems, Learning Problems, Psychosomatic, Impulsive-Hyperactive, and Anxiety. The original version of this scale has been shown to be sensitive to stimulant drug effects (Barkley, 1977). While the revised scale has no research on its drug sensitivity, it was employed here because of its similarity to the original form while requiring far less time by the parents to complete. Although redundant with the information from the CBCL as part of the initial evaluation, the CPRS-R was used in this initial exam to reduce the likelihood that its repeated use in the 3 week drug protocol would result in practice effects that might confound its results. Such effects are known to occur with this instrument and primarily develop between the first and second administrations. It is recommended that it be administered once before using it to assess treatment effects (Barkley, [illegible]).

c. Home Situations Questionnaire (HSQ; Barkley, 1981): This rating scale required parents to answer whether or not a child posed behavior problems in each of 16 situations listed on the scale (i.e., While Playing Alone, When Asked To Do Chores, In Public Places, etc.). If so, the parent indicated the severity of the problem using a scale of 1 (mild) to 9 (severe). Two scores were obtained, these being the total number of problematic settings and the mean severity score. The scale differentiates ADD from normal children (Barkley, 1981; Tarver-Behring, Barkley, & Karlsson,

1985) and is sensitive to stimulant drug effects (Barkley, Karlsson, Pollard, & Murphy, 1985).

5. Parent Self-Report Measures: While the children were completing the laboratory tests, mothers completed the following questionnaires. This was to provide a more comprehensive description of the children's family. Previous research has demonstrated that parenats of ADD children have significantly more problems with stress, depression, and marital discord than those of normal children (Befera & Barkley, 1985).

a. Parenting Stress Index (PSI: Burke & Abidin, 1980): The PSI is a 150 item multiple choice questionnaiore which yields six scores pertaining to child behavioral characteristics (e.g., distractibility, mood, etc.), eight scores pertaining to maternal characteristics (e.g., depression, sense of competence as a parent, etc.), and two scores pertaining to situational and life stress events. These scores can be summed to yield 3 domain or summary scores, these being Child Domain, Mother Domain, and Total Stress.

b. Beck Depression Inventory (Beck, 1967): This is a 21 item multiple choice questionnaire designed to permit a quick assessment of self-reported levels of depression in the respondent. The total score is the sum of the number of credits per item across all items. Higher scores reflect greater depression. The scale was used since prior research has shown that mothers of ADD children may be more depressed than those of normal children (Befera & Barkley, 1985).

c. Locke-Thomes Marital Adjustment Scale (Locke & Thomes, 1980): This is a brief 19 item multiple choice questionnaire designed to permit a quick assessment of marital discord. The score was the sum of the number of credits assigned to each answer across all items. Lower scores

155

reflect greater marital dissatisfaction. Previous research found mothers of ADD children, particularly those of boys, to have higher ratings of discord than mothers of normal children (Befera & Barkley, 1985).

6. Teacher Ratings of Child Behavior: The children's primary teachers were mailed a packet of rating scales to complete and return. This packet included:

a. Child Behavior Checklist - Teacher Report Form (Achenbach & Edelbrock, 1983): This questionnaire has two sections, the first dealing with Adaptive Functioning at School and the second with Behavior Problems. The Adaptive Functioning section can be scored for five scales (School Performance, Working Hard, Behaving Appropriately, Learning, and Happy) or two summary scores, one for School Performance and the second for Total Adaptive Functioning (sum of remaining four scales). The Behavior Problem Scale has 113 items similarly worded to the Parent Report Form described above. Norms are available for two age groups (6 to 11 and 12 to 16 years) for each sex. The scales for 6 to 11 year old boys are: Anxious, Social-Withdrawal, Unpopular, Self-Destructive, Obsessive-Compulsive, Inattentive, Nervous-Overactive, and Aggressive.

b. Conners Teacher Rating Scale - Revised (CTRS-R; Goyette et al., 1978): The CTRS-R is a 28 item questionnaire constructed similarly to the CPRS-R described above. The factors are Conduct Problems, Hyperactivity, and Inattentive-Passive. The original version of this rating scale (Conners, 1969) from which the revised version was constructed is one of the most widely used and sensitive measures of drug responding in ADD children (Barkley, 1977). Again, like the CPRS-R, this scale is redundant with information from the Child Behavior Checklist - Teacher Report Form. However, because it has a significant practice effect between first and second administrations, it was administered as part of the initial evaluation to reduce the likelihood of practice effects on the scale confounding its results when used in the repeated drug assessment battery.

c. School Situations Questionnaire (SSQ; Barkley, 1981): This rating scale is similar to the HSQ except that the settings to be rated deal with situations in the school most likely to be problematic for ADD children. The scale lists 12 situations (i.e., During Individual Desk Work, In the Hallways, During Large Group Work, etc.) and the teacher was asked to indicate whether or not the child was a problem in each. If so, the teacher rated the problem on a scale from 1 (mild) to 9 (severe). Two scores were derived, these being the number of problem settings and their mean severity rating.

Repeated Drug Evaluation Battery. The vast majority of the measures described above were repeated each week during the drug evaluation. Those not repeated were the Child Behavior Checklist (both parent and teacher versions), and the parent self-report measures (PSI, Beck, and Locke-Thomes scales). These measures were not repeated because they have not been shown to be sensitive to stimulant drug effects and were not of interest in the 3 week drug assessment. They were used during the initial evaluation of each child as diagnostic aides or to evaluate various aspects of family functioning pertinent to determining future treatment recommendations for the child or family.

One measure not used in the initial evaluation but repeated during each week of the drug trial was a rating scale of side effects. The Side Effects Rating Scale has been used in previous drug studies to obtain information about the occurrence of 17 possible side effects known to occur with Ritalin (Barkley, 1981). Parents or teachers answered Yes or No as to whether each side effect was noted the previous week and, if so, they rated its severity using a scale from 1 (mild) to 9 (severe). Two scores were derived from both the parent and teacher completed scale: the number of side effects and their mean severity rating. These measures have been shown to be sensitive to dose effects of Ritalin (Barkley et al., 1985).

Results

<u>Subjects</u>

A total of 28 children between 5 and 12 years of age suspected of having ADD were referred to this drug assessment service. There were 21 boys and seven girls. All were believed to have ADD by their referring professionals. Of these 28, three were rejected for reasons described earlier, and two were prematurely discontinued during the evaluation due to tic reactions developing to the medication. The tics subsided within several days after discontinuing the medication in both children. Denckla, Bemporad, and MacKay (1976) found that fewer than 2 percent of a large number of ADD children receiving medication in their clinical practice developed tic reactions. Our results suggest that the reactions may be more common among ADD children exposed to stimulants than had previously been believed. Clinicians should quest in parents thoroughly about such movement disturbances during the drug trials.

To be eligible for the drug evaluation, the children had to meet the following criteria: (1) parent and/or teacher complaints of poor sustained attention, impulsivity, and restlessness; (2) a duration of at least 12 months of these behavior problems; (3) onset of the problems by 6 years of age; (4) scores on the Hyperactivity Factor of either the CPRS-R or CTRS-R of at least 1.5 standard deviations above the mean for same age normal children; (5) no history of mental retardation or significant developmental delay; (6) absence of epilepsy, gross brain damage, gross sensory deficits, or severe emotional disturbance (autism, psychosis, etc.); and (7) have no history of tic disorders or Tourette's Syndrome, cardiovascular problems, or previously poor response to Ritalin after age 5 years.

The initial characteristics of the 23 children (17 males, 6 females) completing the full evaluation and their parents are set forth in Table 1. These findings suggest that this sample was quite deviant from normal with

Place Table 1 about here

respect to parent and teacher ratings of behavior and pervasiveness of the behavior problems across situations. However, some children were identified as ADD (Hyperactivity Factor score greater than 1.5 standard deviations above the normal mean) by their teachers but not their parents on the Conners scales while others were so identified by their parents but not their teachers. Many, however, were rated as ADD on these scales by both informants. The significantly deviant T scores on the Hyperactivity and Inattention factors of both the parent and teachers versions of the CBCL corroborate the success of the screening criteria in selecting a group of significantly inattentive, impulsive, and restless children. A majority of these children appeared to have difficulties with both ADD and conduct problems. Their mothers did not appear to be more disturbed or depressed than normal as indicated by the percentile scores on the PSI, but did report more behavioral problems and stress with their children on this measure. Similarly, the mothers did not rate themselves as more depressed than normal on the Beck Depression Inventory or as having more marital discord than normal on the Locke-Thomes scale, using commonly recommended cutoff scores for deviance on these scales.

On the lab measures, these children had a mean performance which fell in the borderline range (between 5th and 10th percentile) on the Delay Task and

159

made more errors than normal on both comission and omission measures on the Vigilance Task of the GDS. Behavioral observations during the Restricted Academic Situation in the clinic playroom found these children to spend an average of 56 percent of their time engaged in off-task behavior and 35 percent in fidgeting; hardly surprising given the reasons for their referral but validating the success of this observational method at detecting these problems. The children engaged in vocal noises or self-directed speech 30 percent of the time, and spoke to their mothers against restrictions not to do so approximately 25 percent of the time. The children engaged in negative behaviors such as whining, refusal to do work, or tantrums approximately 7 percent of the time and their mothers initiated commands toward the children during 14 percent of the observations. Overall, ADD behaviors were noted during an average of 25 percent of the observation intervals.

Reliability of Measures

When utilizing measures in a repeated assessment battery such as this, it is essential that the measures have satisfactory test-retest reliability between repeated administrations where no treatment occurred in the interim. Several measures have such information available in the literature or from their developers. Gordon (personal communication) has reported test-retest coefficients for 90 normal children over a 30 to 45 day interval for the Delay Task of .77 for Total Responses, .68 for Number Correct, and .60 for Efficiency Ratio. For the Vigilance Task, coefficients were .85 for Total Correct, .72 for Comission Errors, and .80 for Omission Errors. For the CTRS-R, Edelbrock, Greenbaum, and Conover (1985) recently reported 1 week test-retest coefficients of .95 for Conduct Problems, .95 for Hyperactivity, .88 for Inattention, and .96 for the Hyperactivity Index. All coefficients were statistically significant in these studies.

160

In order to gain information about the test-retest reliability on the remaining measures, scores from the initial assessment were correlated with those taken during the placebo condition. Such a method will likely yield lower estimates of reliability than the more traditional method of using fixed time intervals between two administrations of a measure in that both the number of weeks (1 to 3) and hence the number of repeat administrations of the measures varied across subjects (e.g., some children received placebo the first week, others the second, and others the third week).

For the CPRS-R, coefficients across this 1 to 3 week interval were .65 for Conduct Problems, .57 for Impulsive-Hyperactive, and .53 for the Hyperactivity Index. For the HSQ, reliability was .66 for the Number of Problem Settings and .62 for the Mean Severity Rating. Coefficients for the same measures for the SSQ were .78 and .63, respectively. For the Total ADD Behaviors taken in the Restricted Academic Situation, reliability was .86. All coefficients were statistically significant. Considering that these are lower bound estimates of reliability, the actual test-retest reliability of these measures using more traditional procedures would certainly be higher. Hence, the reliability of all measures seems quite adequate. No reliability was calculated for the Side Effects Rating Scale as it was not administered as part of the initial evaluation of the children but only during the repeated drug evaluations.

Drug Effects

All dependent measures used in the repeated assessment battery across the drug conditions were submitted to one-way (drug condition) analyses of variance with repeated measures. The means, standard deviations, and results from the statistical tests are set forth in Table 2. Significant main

161

Place Table 2 about here

effects for drug condition were noted on 16 of the 31 dependent measures. The majority of these effects were found on the teacher ratings and the behavioral observations taken in the clinic during the Restrictedd Academic Situation. Surprisingly, significant drug effects were not noted on the percent of math problems performed correctly during this playroom observation, but this may have resulted from the use of different types of math problems across the three testing occasions. The number and rated severity of side effects were also not significantly affected by these two doses of medication. In view of the many drug related improvements in behavior, this suggests that the dose levels were well within a judicious therapeutic range for these children.

Pairwise comparisons were conducted on the 16 measures having significant main effects and these results are shown in Table 3. On the vast majority of these measures, the low and

Place Table 3 about here

high doses were equally effective in improving the behavior of these children. Only on the behavioral category of "Plays with Objects" during the Restricted Academic Situation was a significant dose effect noted for the high dose compared to the low dose. While both doses resulted in a significant reduction in this behavior, the higher dose proved more effective. Although significant drug effects were noted on the Vigilance Task of the GDS, only the high dose of Ritalin resulted in significant improvements on both Comission and Omission Errors.

162

At the end of each child's evaluation, the results were reviewed by the psychologist supervising the drug trial and a recommendation was made to the referring physician as to which dose level, if any, appeared to have been most effective for this child. All results entered into this clinical decision with that dose making the greatest changes in behavior across the most measures and with the least side effects being chosen as the best dose for the child at that time. These clinical judgements resulted in 5 (20%) of the 25 children entering the protocol not recommended for any medication (2 of these were the children discontinuing prematurely due to tic reactions), 6 (24%) recommended for the low dose, 8 (32%) recommended for the higher dose, and 6 (24%) having a moderate dose between the high and low doses recommended. This latter recommendation was made where it was clear that both doses resulted in an effective drug response but the higher dose resulted in greater side effects for that particular child. Hence, a total of 80% of the children entering the protocol were recommended for at least some level of Ritalin following this drug-placebo evaluation. Comparisons on the initial evaluation measures of the responders and nonresponders were not undertaken because of the inadequate sample sizes per group to permit satisfactory statistical power in such comparisons.

Discussion

The results of this drug study suggest that an easily administered, cost-effective, multi-method assessment battery can be quite useful in the routine clinical evaluation of stimulant drug responding in ADD children. Such a test battery draws upon multiple sources and types of information using several methods of assessment and yields a wealth of clinically useful information on each case. The battery can be conducted in less than 1 hour per week, can be accomplished for a reasonable clinic fee (net cost for the 5

163

hours of evaluation for personnel, resources, drugs, and supplies was $120), and can be easily conducted by a paraprofessional with a short period of prior training in administering the tests (further enhancing its cost effectiveness).

The most useful measures for detecting stimulant drug effects were those derived from the teacher ratings and from the behavioral observations of academic performance in the clinic playroom. While some may be tempted to eliminate the use of the direct observational measures in favor of using only the more easily obtained teacher ratings, we recommend against it. First, the playroom measures proved equally if not more sensitive to the drug effects than did the teacher ratings of behavior in school. Second, it was our experience that such behavioral observations lend added credibility to the assessment protocol from the view of parents, teachers, and referring professionals in that they provided an objective assessment of the child's behavior apart from parent and teacher opinions. Third, when teacher ratings may not always be so readily obtained, in cases of poor teacher cooperation, during summer vacations, or when teacher ratings are believed to be unreliable, the clinic playroom measures are a reasonable substitute. Finally, it is useful in such assessments to build in a modicum of redundancy across measures in order to corroborate the drug effects noted on one method taken in one setting from one source against those from a different method, setting, and source.

Similarly, while it would seem at first glance more efficient to dispense with the parent ratings since they were of little value in detecting drug or dose effects, this, too, would be a mistake. These measures were useful in the initial diagnosis of the children and provided a necessary vehicle for parents to inform clinic staff of their observations across drug

164

conditions. In some cases, these ratings revealed significant drug effects or side effects of use in further treatment planning. In fact, several parent ratings (hyperactive-impulsive behavior and mean severity score of problem behaviors) did show significant drug effects, particularly at the higher dose. Moreover, such ratings were valuable in providing ecologically valid assessments of behavior in an important setting (home) as reported by a very important observer (parents), whose views in part resulted in the initial referral of the child. Nor should it be at all surprising that drug effects were limited on the parent reported data as parents have much less opportunity to observe drug effects given their short time course (6 to 7 hours) and that such effects have often dissipated shortly after the child has returned home from school.

The failure of the GDS Delay Task to prove sensitive to drug effects was surprising in view of its purported evaluation of impulsivity -- a behavior typically responsive to stimulant drug effects -- and its previously demonstrated ability to differentiate ADD from non-ADD clinic-referred children (Gordon, 1979). Perhaps practice effects develop on this measure over repeated administrations (Douglas, personal communication) such that a ceiling effect develops beyond which medication effects cannot be detected. This would suggest that varying the delay interval across the three repeated assessments and forewarning the child in each session that the waiting time between button presses is different from that of the previous week might help to overcome this effect. The first author (RAB) is currently testing this procedural variation in an ongoing drug study. The Kagan Matching Familiar Figures Test (Kagan, 1966), a commonly used measure of impulsivity, is also being studied for comparison purposes because of its previously demonstrated sensitivity to drug effects (Barkley, 1977).

Significant drug effects were noted on the GDS Vigilance Task for both Comission and Omission Errors, but only at the high dose of Ritalin. This is difficult to explain considering that vigilance tasks of this sort are generally quite sensitive to lower doses of medication in testing periods even shorter than that used here (Sykes, Douglas, Weiss, & Minde, 1971; Werry & Aman, 1975). Perhaps extending the length of this test, say from 9 to 12 minutes, might enhance its sensitivity to the lower doses of medication. Nevertheless, retention of this measure in the drug protocol is recommended for the time being as it remains one of the few objective measures of inattention in the protocol aside from the direct observational measures and it has been shown to be dose sensitive in another drug study reported in this same journal issue.

The addition of a brief of verbal learning and memory task to this protocol might enhance the sensitivity of the laboratory tasks to stimulant drug effects while providing a measure of drug effects on a different domain of cognitive ability not presently assessed yet of import to academic performance. Numerous prior stimulant drug studies have found the paired associate learning task (PAL) to be sensitive to varying doses of stimulant medication (Dalby et al., 1977; Douglas, Barr, O'Neill, & Britton, in press; Rapport, Stoner, DuPaul, Birmingham, & Tucker, 1985). The particular version of this task used by Douglas et al. (in press) appears to be the easiest and least time consuming to administer, as well as the least expensive of these tasks.

The results of the present study argue strongly for the development of regional drug evaluation clinics to which area professionals can refer ADD children for objective, multi-method yet economical stimulant drug evaluations to replace the currently less reliable and more subjective

166

approach employed in routine clinical practice. Many medical schools, hospitals, or mental health centers have the personnel, financial, and physical resources available to easily conduct these protocols and rapidly return useful information to the referring professional for the ongoing care and management of ADD children. Moreover, the use of a relatively consistent protocol such as this one across numerous regional drug assessment centers which combine their findings could also greatly accelerate the collection of scientific information on much larger subject samples than has heretofore been the case. In any event, the relatively high prevalence of the disorder and the widespread use of stimulant drugs to treat it demand no less than the most objective, reliable, valid, and cost-effective routine clinical assessment protocol our present scientific knowledge and technology can justify.

References

Achenbach, T. M., & Edelbrock, C. (1983). Manual for the Child Behavior Checklist and Revised Child Behavior Profile. Burlington, VT: Thomas Achenbach.

Barkley, R. A. (1977). A review of stimulant drug research with hyperacative children. Journal of Child Psychology and Psychiatry, 18, 137-165.

Barkley, R. A. (1981). Hyperactive children: a handbook for diagnosis and treatment. New York: Guilford Press.

Barkley, R. A. (1986a). Attention deficit disorders. In E. Mash & L. Terdal (Eds.) Behavioral assessment of childhood disorders (second edition). New York: Guilford.

Barkley, R. A. (1986b) A review of child behavior rating scales and checklists for research in child psychopathology. In M. Rutter (Ed.) Behavioral assessment methods for research in child psychopathology. New York: Guilford.

Barkley, R. A., Karlsson, J., Pollard, S., & Murphy, J. (1985). Developmental changes in the mother-child interactions of hyperactive boys: effects of two doses of Ritalin. Journal of Child Psychology and Psychiatry, 26, 705-715.

Beck, A. T. (1967). Depression: causes and treatment. Philadelphia, PA: University of Pennsylvania Press.

Befera, M., & Barkley, R. A. (1985). Hyperactive and normal girls and boys: mother-child interactions, parent psychiatric status, and child psychopathology. Journal of Child Psychology and Psychiatry, in press.

Burke, W. T., & Abidin, R. R. (1980). Parenting Stress Index (PSI): a family system assessment approach. In R. R. Abidin (Ed.) Parent education and intervention handbook. Springfield, IL: Charles C. Thomas.

Cantwell, D. (1980). A clinician's guide to the use of stimulant medication for psychiatric disorders of children. Developmental and Behavioral Pediatrics, 1, 133-140.

Cantwell, D., & Carlson, G. (1978). Stimulants. In J. Werry (Ed.) Pediatric psychopharmacology. New York: Bruner/Mazel.

Comings, D. E., & Comings, B. G. (1984). Tourette's Syndrome and Attention Deficit Disorder with Hyperactivity: Are they genetically related? Journal of the American Academy of Child Psychiatry, 23, 138-146.

Conners, C. K. (1969). A teacher rating scale for use in drug studies with children. American Journal of Psychiatry, 126, 152-156.

Dalby, J. T., Kinsbourne, M., Swanson, J. M., & Sobol, M. P. (1977). Hyperactive children's underuse of learning time: correction by stimulant treatment. Child Development, 48, 1448-1453.

Denckla, M. B., Bemporad, J. R., & MacKay, M. C. (1976). Tics following methylphenidate administration: a report of 20 cases. Journal of the American Medical Association, 235, 1349-1351.

Douglas, V. I. (1985). Personal communication, October, 1985.

Douglas, V. I., Barr, R. G., O'Neill, M. E., & Britton, B. G.

(1986). Short term effects of methylphenidate on the cognitive, learning, and academic performance of children with Attention Deficit Disorder. Journal of Child Psychology and Psychiatry, in press.

Dulcan, M. K. (1985). The psychopharmacologic treatment of children and adolescents with Attention Deficit Disorder. Psychiatric Annals, 15, 69-87.

Edelbrock, C., Costello, A., & Kessler, M. D. (1984). Empirical corroboration of Attention Deficit Disorder. Journal of the American Academy of Child Psychiatry, 23, 285-290.

Golden, G. S. (1982). Movement disorders in children. Developmental and Behavioral Pediatrics, 3, 209-216.

Gordon, M. (1979). The assessment of impulsivity and mediating behaviors in hyperactive and non-hyperactive children. Journal of Abnormal Child Psychology, 7, 317-326.

Gordon, M. (1983). The Gordon Diagnostic System. Boulder, CO: Clinical Diagnostic Systems.

Gordon, M. (1986). Personal communication, January, 1986.

Goyette, C. H., Conners, C. K., & Ulrich, R. F. (1978). Normative data for Revised Conners Parent and Teacher Rating Scales. Journal of Abnormal Child Psychology, 6, 221-236.

Hastings, J. E., & Barkley, R. A. (1978). A review of psychophysiological research with hyperkinetic children. Journal of Abnormal Child Psychology, 6, 413-447.

Kagan, J. (1966). Reflection-impulsivity: the generality and dynamics of conceptual tempo. Journal of Abnormal Psychology, 71, 17-24.

Klee, S. H., & Garfinkel, B. D. (1983). The computerized continuous performance task: a new measure of inattention. Journal of Abnormal Child Psychology, 11, 487-496.

Locke, H. J., & Thomes, M. M. (1980). The Locke Marital Adjustment Test: its validity, reliability, weighting procedure, and modification. Unpublished manuscript, University of Southern California.

Mash, E., & Terdal, L. (Eds.). (1981). Benavioral assessment of childhood disorders. New York: Guilford.

Milich, R. (1984). Cross-sectional and longitudinal observations of activity level and sustained attention in a normative sample. Journal of Abnormal Child Psychology, 12, 261-276.

Milich, R., Loney, J., & Landau, S. (1982). The independent dimensions of hyperactivity and aggression: a validation with playroom observation data. Journal of Abnormal Psychology, 91, 183-187.

Milich, R., Loney, J., & Whitten, P. (1980). Two year stability and validity of playroom observations of hyperactivity. Paper presented at the annual meeating of the American Psychological Association, Anaheim, CA.

Porteus, S. D. (1965). Porteus maze tests: fifty years of application. Palo Alto, CA: Pacific Books.

Rapport, M. D., Stoner, G., DuPaul, G. J., Birmingham, B. K., & Tucker, S. (1985). Methylphenidate in hyperactive children: differential effects of dose on academic, learning, and social behavior. Journal of Abnormal Child

Psychology, *13*, 227-244.

Ross, D., & Ross, S. (1976). *Hyperactivity*. New York, Wiley.

Ross, D., & Ross, S. (1982) *Hyperactivity* (second edition). New York: Wiley.

Sykes, D. H., Douglas, V. I., Weiss, G., & Minde, K. K. (1971). Attention in hyperactive children and the effect of methylphenidate (Ritalin). *Journal of Child Psychology and Psychiatry*, *12*, 129-139.

Tarver-Behring, S., Barkley, R., & Karlsson, J. (1985) The mother-child interactions of hyperactive boys and their normal siblings. *American Journal of Orthopsychiatry*, *55*, 202-209.

Ullman, D. G., Egan, D., Fiedler, N., Jurenac, G., Pliske, R., Thompson, P., & Doherty, M. E. (1981). The many faces of hyperactivity: similarities and differences in diagnostic policies. *Journal of Consulting and Clinical Psychology*, *49*, 694-704.

Varley, C. K., & Trupin, E. W. (1983). Double-blind assessment of stimulant medication for Attention Deficit Disorder: a model for clinical application. *American Journal of Orthopsychiatry*, *53*, 542-547.

Werry, J. S., & Aman, M. (1975). Methylphenidate and haloperidol in children: effects on attention, memory, and activity. *Archives of General Psychiatry*, *32*, 790-794.

Footnotes

1. Portions of this paper were presented at the meeting of the American Psychological Association, Los Angeles, CA, August, 1985. The authors wish to thank Kathy Schoendorf for her assistance with the data collection and transcription and Lewis Bloomingdale, M.D., for his comments on previous drafts of this manuscript.

Table 1

Initial Subject Characteristics

Measure	Mean	SD*	Range
Age (in years)	8.5	2.3	5.[illegible] - [illegible]2.9
Child's Education (in years)	2.9	2.1	0 - 6
Age of Onset of ADD (in years)	4.6	1.5	1 - 6.9
Mother's Age (in years)	34.1	5.6	27 - 44
Mother's Education (in years)	13.3	1.9	11 - 16
Home Situations Questionnaire:			
Number of Problem Settings	9.1	4.0	2 - 16
Mean Severity of Problems	4.3	1.7	1 - 6.8
Conners Parent Rating Scale:			
Conduct Problems	14.7	7.3	3 - 27
Hyperactive-Impulsive	8.4	3.0	0 - 12
Hyperactivity Index	18.5	5.0	5 - 26
Child Behavior Checklist: (T scores)			
Hyperactivity	72.2	7.8	56 - 85
Delinquency	66.6	7.8	56 - 86
Aggression	71.2	10.8	55 - 89
School Situations Questionnaire:			
No. of Problems	7.4	2.8	2 - 12
Mean Severity	5.4	1.4	2.7 - 7.5
Conners Teacher Rating Scale:			
Conduct Problems	9.1	6.1	6 - 23
Hyperactivity	13.1	5.9	2 - 21

Inattention	14.4	5.5	6 - 24
Hyperactivity Index	17.9	5.4	8 - 26
Child Behavior Checklist - Teacher: (T scores)			
Inattention	70.2	8.4	56 - 87
Nervous-Overactive	69.1	7.9	55 - 86
Aggression	66.4	7.3	55 - 83
Beck Depression Scale	7.5	6.2	1 - 27
Locke-Thomas Marital Adjustment Scale	114.9	31.5	38 - 160
Parent Stress Index (Percentiles):			
Total Stress Domain	72.0	23.9	30 - 100
Child Domain	91.5	10.6	65 - 100
Mother Domain	54.8	31.1	10 - 95
Parent Depression Scale	54.1	29.6	10 - 95

* SD = standard deviation

Table 2

Means, Standard Deviations, and Statistical Test Results on the Dependent Measures for Each Drug Condition

Measure	Placebo Mean	Placebo SD	Low Dose Mean	Low Dose SD	High Dose Mean	High Dose SD	F	p
Home Situations Questionnaire:								
No. of Problems	8.0	4.5	7.4	4.2	7.3	5.1	0.54	-
Mean Severity	4.1	1.9	3.2	1.8	3.0	1.8	4.00	.05
Conners Parent Scale:								
Conduct Problem	10.7	7.7	8.8	6.4	9.3	6.0	0.87	-
Hyperactive	7.0	3.3	5.0	3.2	5.1	3.4	7.19	.01
Index	13.3	6.9	10.8	6.1	10.7	7.4	2.20	-
Side Effects - Parent:								
Number	5.0	4.0	5.1	3.9	5.2	3.3	0.08	-
Mean Severity	3.4	1.7	2.6	1.7	3.0	1.4	2.27	-
School Situations Questionnaire:								
No. of Problems	5.8	3.3	4.5	2.7	5.0	3.4	4.32	.05
Mean Severity	4.3	2.3	2.6	1.8	2.6	1.9	11.03	.01
Conners Teacher Scale:								
Conduct Problem	7.2	6.0	4.2	4.4	3.9	3.8	6.50	.01
Hyperactivity	11.1	6.5	6.7	5.1	6.5	5.4	7.97	.01
Inattention	10.8	5.0	8.0	4.7	7.6	5.1	7.60	.01
Index	14.3	7.1	8.9	5.7	8.7	5.4	10.69	.01
Side Effects - Teacher:								
Number	3.6	3.3	3.3	2.8	4.1	3.8	0.92	-
Mean Severity	3.0	2.3	2.8	2.1	2.5	2.3	0.34	-

Stimulants and ADD

GDS - Delay Task:								
No. of Rewards	40.1	13.0	41.4	19.0	41.7	15.2	0.19	-
No. of Responses	71.5	42.2	74.5	61.1	66.7	35.6	0.58	-
Efficiency Ratio	66.8	27.4	73.7	28.7	74.5	28.6	2.54	-
GDS - Vigilance:								
No. Correct	40.8	18.9	39.3	15.6	46.2	14.9	2.09	-
Commissions	14.5	19.4	16.7	22.4	6.1	[illegible]	3.41	.05
Omissions	10.0	12.2	9.6	15.5	4.5	7.1	3.24	.05
Playroom Observations (percent):								
Off-task	56.1	32.3	45.4	25.7	41.5	24.8	5.45	.01
Fidgets	38.3	25.4	27.4	20.8	30.4	23.3	2.10	-
Vocalizes	38.9	32.9	29.0	27.5	30.5	30.4	2.06	-
Talks to Mother	26.5	33.3	19.3	26.5	14.5	23.0	5.42	.01
Plays w/ Obj.	14.8	20.3	7.2	13.4	1.7	4.0	8.83	.01
Out-of-seat	12.6	24.7	5.4	16.4	2.5	6.7	5.82	.01
Child Negative	5.6	9.8	6.1	12.6	1.9	7.0	1.93	-
Mother Commands	15.8	25.6	9.7	19.4	7.2	13.9	3.35	.05
Total ADD Beh.	27.1	17.8	18.7	12.1	16.3	9.8	12.60	.01
Percent Correct in								
Math Problems	70.0	31.9	74.5	31.9	76.8	22.5	1.24	-

SD = standard deviation

F = F ratio from analysis of variance

p = probability value of F-test from analysis of variance

Table 3

Probability Values From Statistical Tests of Pairwise Comparisons on Measures Having Significant Main Effects

Measure	P vs L	P vs H	L vs H
Home Situations Questionnaire:			
Mean Severity	-	.01	-
Conners Parent Scale:			
Hyperactive/Impulsive	.01	.01	-
School Situations Questionnaire:			
No. of Problems	.05	-	-
Mean Severity	.01	.01	-
Conners Teacher Scale:			
Hyperactivity	.01	.01	-
Inattentive	.01	.01	-
Conduct Problems	.05	.01	-
Hyperactivity Index	.01	.01	-
GDS - Vigilance Task:			
No. of Commissions	-	.05	.05
No. of Omissions	-	.05	.07
Playroom Observations:			
Off-task	.05	.01	-
Talks to Mother	.05	.05	-
Plays with Objects	.05	.01	.05
Out-of-seat	.01	.05	-
Mother Commands	.05	.05	-
Total ADD Behaviors	.01	.01	-

Stimulants and ADD

P = placebo condition

L = low dose condition

H = high dose condition

STANDARDIZATION OF THE GORDON DIAGNOSTIC SYSTEM

The Gordon Diagnostic System (GDS) was standardized on 1356 children. This current sample is comprised of 716 boys and 640 girls 3 to 16 years of age. There are 306 3-5 year olds, 337 6-7 year olds, 263 8-9 year olds, 221 10-11 year olds and 229 12-16 year olds.

The children were randomly selected from class lists from both public and private schools. According to the Hollingshead measure of socioeconomic status, subjects came from the full range of social classes. The children were primarily from the greater Syracuse area as well as from Virginia and Philadelphia. The children were administered the GDS on an individual basis in school by trained examiners. A child's scores were included in the standardization sample if the child had no known psychiatric problems, learning problems, had not been retained, had never received psychoactive medication, and showed no evidence of frank neurological damage.

GDS scores have been found unaffected by SES or geographic location. Comparison with the norms of an inner-city sample of impoverished children and a sample from an impoverished rural area revealed no significant differences in mean scores.

There currently is a standardization study being conducted by Dr. Jose Bauermeister with Puerto Rican children. The sample contains approximately 200 children who were randomly selected from class lists of public and private schools. Another large project is being conducted with Department of Defense dependents in Okinawa by Sandra Sacco.

Analyses of standardization data have revealed no significant sex differences. Consequently, identical norms are presented for boys and girls.

As indicated in the GDS Interpretive Supplement (page 1), the Threshold Tables divide the standardization sample into three groups according to statistical convention in the behavioral sciences. As such, a child's score is considered ABNORMAL if it falls in the 5th percentile or lower. The BORDERLINE range includes children with scores in the 6th - 25th percentile, and the NORMAL range is for those with scores above the 25th percentile.

INFORMATION

REGARDING

STANDARDIZATION

WIDE RANGE ACHIEVEMENT TEST-REVISED

AGE NORMS

LEVEL 1

AGE: 11-6 to 11-11

Stand. Score	RAW SCORE Reading	RAW SCORE Spelling	RAW SCORE Arithmetic	Percentile
46	43	—	—	.03
47	44	24	—	.04
48	—	—	16	.05
49	45	25	—	.06
50	[illegible]	[illegible]	[illegible]	.07
51	46	—	17	.08
52	47	26	—	.09
53	—	—	—	.1
54	48	27	18	.2
55	49	—	—	.3
56	—	28	—	.4
57	50	—	19	.5
58	—	29	—	.6
59	51	—	—	.7
60	52	30	20	.8
61	—	—	—	.9
62	53	31	—	1
63	54	—	21	1
64	—	—	—	1
65	55	32	—	1
66	—	—	22	1
67	56	33	—	1
68	57	—	—	2
69	—	34	23	2
70	58	—	—	2
71	—	35	—	3
72	59	—	—	3
73	60	36	24	4
74	—	—	—	4
75	61	37	—	5
76	62	—	25	5
77	—	—	—	6
78	63	38	—	7
79	—	—	26	8
80	64	39	—	9
81	65	—	—	10
82	—	40	27	12
83	66	—	—	13
84	67	41	—	14
85	—	—	28	16
86	68	42	—	18
87	—	—	—	19
88	69	43	29	21
89	70	—	—	23
90	—	—	—	25
91	71	44	30	27
92	—	—	—	30
93	72	45	—	32
94	73	—	31	34
95	—	46	—	37
96	74	—	—	39
97	75	47	—	42
98	—	—	32	45
99	76	48	—	47
100	—	—	—	50

Stand. Score	RAW SCORE Reading	RAW SCORE Spelling	RAW SCORE Arithmetic	Percentile
101	77	49	33	53
102	78	—	—	55
103	—	—	—	58
104	79	50	34	61
105	80	[illegible]	[illegible]	63
106	—	51	—	66
107	81	—	35	68
108	—	52	—	70
109	82	—	—	73
110	83	53	36	75
111	—	—	—	77
112	84	54	—	79
113	85	—	37	81
114	—	55	—	82
115	86	[illegible]	[illegible]	84
116	—	—	38	86
117	87	56	—	87
118	88	—	—	88
119	—	57	39	90
120	89	—	[illegible]	91
121	—	58	—	92
122	90	—	—	93
123	91	59	40	94
124	—	—	—	95
125	92	60	—	95
126	93	—	41	96
127	—	61	—	96
128	94	—	—	97
129	—	—	42	97
130	95	62	—	98
131	96	—	—	98
132	—	63	43	98
133	97	—	—	99
134	98	64	—	99
135	—	—	44	99
136	99	65	—	99
137	—	—	—	99
138	100	—	45	99
139	—	—	—	99.1
140	—	—	—	99.2
141	—	—	46	99.3
142	—	—	—	99.4
143	—	—	—	99.5
144	—	—	47	99.6
145	—	—	—	99.7
146	—	—	—	99.8
147	—	—	—	99.9
148	—	—	48	99.91
149	—	—	—	99.92
150	—	—	—	99.93
151	—	—	49	99.94
152	—	—	—	99.95
153	—	—	—	99.96
154	—	—	50	99.97
155	—	—	—	99.98

181

THE GORDON DIAGNOSTIC SYSTEM

GDS RECORD FORM

Name: .... Date Tested: ....

Sex: .... Date of Birth: ....

Age: ....

Referred by: ....

Presenting Problem: Duration: ....

School: .... Grade: ....

Special Academic Programming:

Remarks:

CLINICAL DIAGNOSTICS INC
1536 Cole Blvd., Golden, Colorado 80401
Suite 350 • 303-232-7329

182

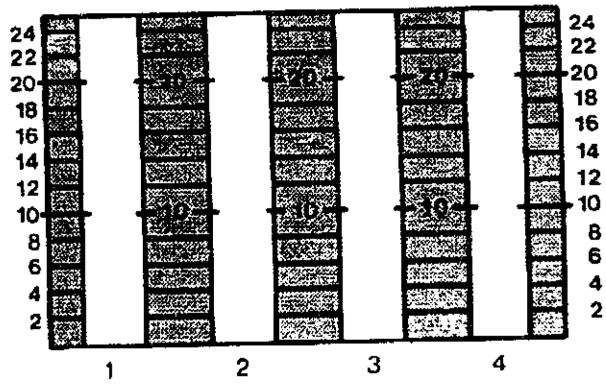
GDS RECORD FORM

DELAY TASK

TASK PARAMETERS

☐ Standard	Delay Interval = 6 Seconds Block Length = 120 Seconds	
☐ Other Fixed Parameters	Delay Interval .... Seconds Block Length .... Seconds	
☐ Variable Parameters	Block 1 2 3 4 Delay Interval .... Block Length	

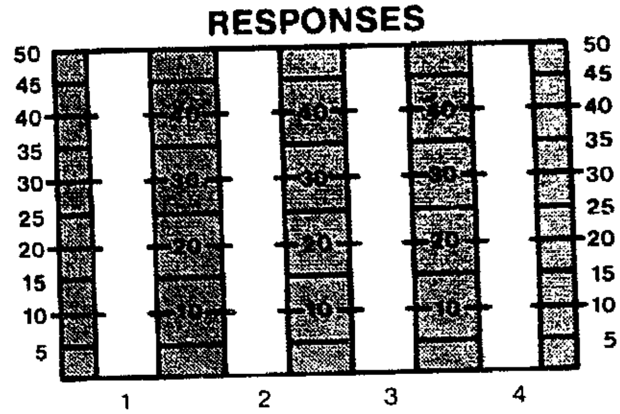
REWARDS



RESULTS

Data Selector Positions	Display		
1	....	Total:	Rewards
2	....		Responses
3	....	Block 1:	Rewards
4	....		Responses
5	....	Block 2:	Rewards
6	....		Responses
7	....	Block 3:	Rewards
8	....		Responses
9	....	Block 4:	Rewards
10	....		Responses

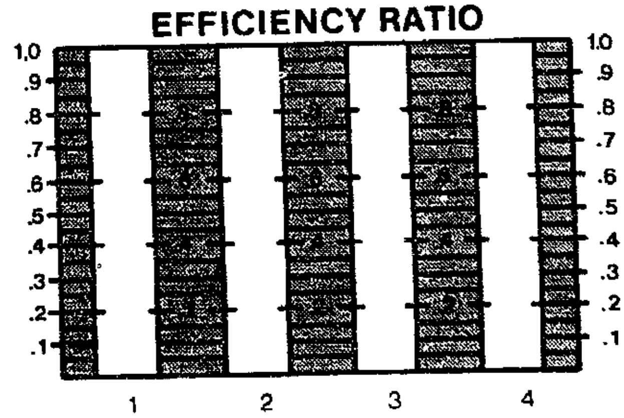
RESPONSES



SUMMARY STATISTICS

Block	1	2	3	4	Total
Interval					
Length					
Rewards					
Responses					
Effic. Ratio					

EFFICIENCY RATIO



STRATEGIES/COMMENTS:

183

GDS RECORD FORM

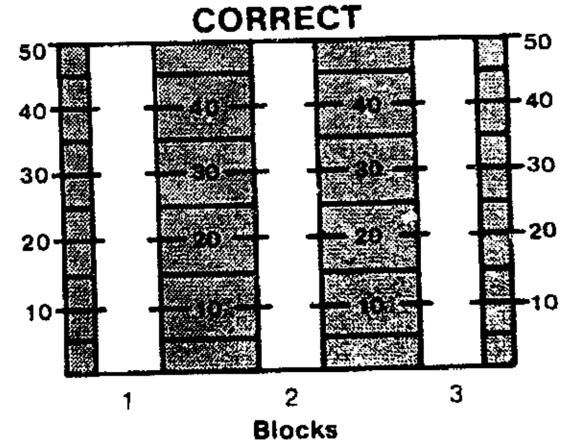
VIGILANCE TASK

TASK PARAMETERS

- ☐ Standard — Pres. Interval = 1 Second; Block Length = 180 Seconds
- ☐ Other Fixed Parameters — Pres. Interval .... Seconds; Block Length .... Seconds
- ☐ Variable Parameters

Block	1	2	3
Presentation Interval	....	....	....
Block Length	....	....	....

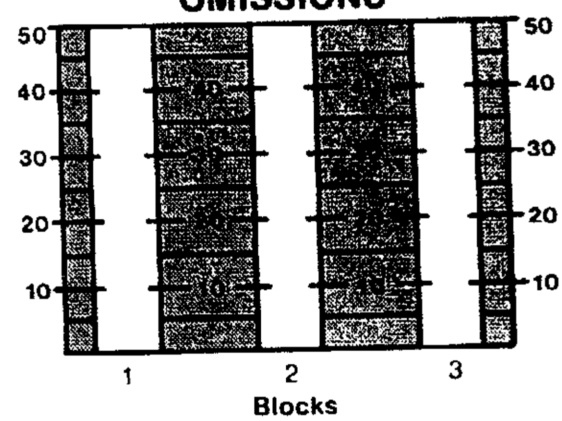
CORRECT



RESULTS

Data Selector Positions	Display		
1	....	Total:	Correct
2	....	Block 1:	Correct
3	....		Omissions
4	....		Commissions
5	....	Block 2:	Correct
6	....		Omissions
7	....		Commissions
8	....	Block 3:	Correct
9	....		Omissions
10	....		Commissions

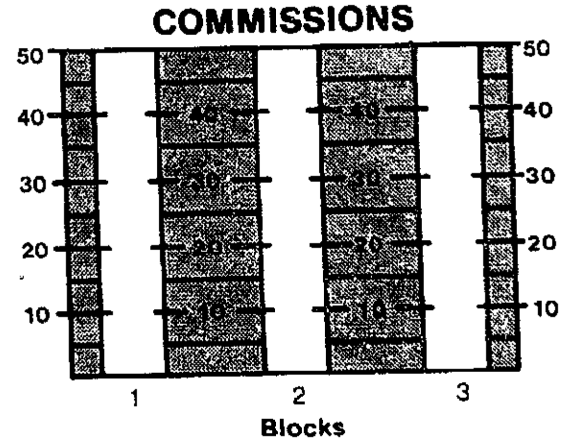
OMISSIONS



SUMMARY STATISTICS

Block	1	2	3	Total
Interval				
Length				
Correct				
Omissions				
Commissions				

COMMISSIONS



STRATEGIES/COMMENTS:

GDS RECORD FORM

ADDITIONAL DATA

WISC-R

Full Scale IQ ....

Verbal IQ ....

Performance IQ ....

Freedom from Distractibility Factor ....

RATING SCALES

Connors Hyperactivity Factor ....

CBCL Hyperactivity Factor ....

Other: ....

BENDER

Koppitz Score ....

Age Equivalent ....

Emotional Indicators ....

Significant Test Behaviors (Neurogenic Compensations) ....

OTHER SOURCES OF INFORMATION

Classroom Observations:

Teacher Reports/Physician Reports:

Medical, Social or Developmental Reports:

DIAGNOSTIC IMPRESSIONS:

185

FURTHER BACKGROUND ON VALIDATION OF GDS

Current GDS Research:

The Vicissitudes of Validation

Michael Gordon, Ph.D.

SUNY/Upstate Medical Center

Paper presented at a symposium entitled, "The Objective Assessment of ADD/Hyperactivity: Research on the Gordon Diagnostic System". Symposium presented at the 93rd Annual Convention of the American Psychological Association at Los Angles, California, August 1985.

186

My original intention for this presentation was to dazzle you with immense quantities of data about the Gordon Diagnostic System (GDS; Gordon, 1982). As it happens, we have just completed two major projects involving more than 800 subjects and a wide range of meaures. But in reviewing these data and considering time limitations, I felt that it would be worthwhile to err more on the side of selectivity than expansiveness in the material I presented. What I would like to do is use these data to g beyond the GDS, per se, and to illustrate what I see as several critical issues surrounding the evaluation of ADD/Hyperactivity.

All of us involved in the diagnosis and treatment of ADD/Hyperactivity are painfully aware of its complexities. This is a disorder comprised of multiple diagnostic dimensions, information about which is often based almost entirely upon the perceptions of different raters with varying internal norms, or upon objective measures which target differential aspects of functioning. But while we all acknowlege the complexity, most of us, myself very much included, tend to shy away from it. We talk more than we should about hyperactive children as if they were a relatively homogenous group which could be evaluated using agreed-upon criteria. We speak of one rating scale being better than the other in selecting the hyperactive child, or about which medications work and which do not, or about how this lab measure misses in the identification of hyperactive children while this one does not -- all with more enthusiasm than may be warranted.

Many of the complexities which need to be respected are illustrated in some of the data I will present to you now. If nothing else, they are intended to highlight how careful we need to be in approaching this diagnostic problem. They also dramatize how seemingly small changes in criteria for group definition can lead to significant changes in results.

What I plan to do is present to you the statistical trail of just one GDS-generated score -- that is, Total Commissions. My hope is that you will put yourself in the shoes of a test developer for a few minutes and confront the following problem: How do you best go about validating a diagnostic procedure for a disorder for which there are no satisfactory or agreed-upon benchmark criteria and where the entity being studied is multidimensional?

The Vigilance Task is a version ofthe venerable Continuous Performance Test (Rosvold et al., 1956). The child sits in front of the GDS and is told that numbers are going to flash on the screen one at a time. They are asked to press the butt[illegible]n only when they see a "1" that comes right after a "9". The GDS records how many times they press the button correctly, how many hot "1/9" combinations they missed, and how many times they pressed the button when they weren't suppose to. This last score represents the Errors of Commission. Typically, they reflect the number of times, following appearance of a "1", the child pressed the button on a digit other than a "9". In other words, the child responds to any number following the "1" without waiting a few extra milliseconds to make sure it is a "9".

Now, one way to establish the utility of a score is to take what might be called the "Arrogant Face Validity" approach. Behind this strategy is the sentiment, "If you can't show me benchmark criteria against which I can establish the accuracy of my test, I'm going to go ahead and assume that the test does indeed measure a salient feature of the disorder. After all, children who don't attend well or who are impulsive when required to attend are bound to do poorly on this test. And besides, there's 25 years of research demonstrating the basic validity of this technique. So I am going to standardize it to the hilt and at least I'll be able to determine with some assurance the degree to which a particular score is normal or abnormal. If nothing else, I'll have some data based on a child's actual behavior".

This "Arrogant Face Validity" approach is one we have pursued intently over the past three years. We have gathered protocols on over 1000 nonhyperactive school children. It was important to me that we standardize the GDS on a nonhyperactive school sample as opposed to the more typical nonhyperactive clinic sample. As a clinician, I am more interested in how a child's performance compares to that of peers than I am in how they compare to the scores of children referred for a range of psychiatric problems. Parenthetically, the insistence that we gather data on a large nonpsychiatric population is one reason why the GDS is not microcomputer-based. To conduct large N studies, I needed an instrument that was light, portable, and absolutely reliable in its administration of the tasks. I also did not want to worry about the child being able to invade disk drives and adjustment buttons on monitors, or about variations in the pressures necessary to actuate a joystick.

Presented here are standardization data for the Total Commission score for children 3 - 16 years of age:

Standardization Table for Total Commissions

AGE	Normal	Borderline	Abnormal
6-7 (n=100)	0 - 10	11 - 23	$\geq$ 24
8-9 (n=96)	0 - 8	9 - 21	$\geq$ 22
10-11 (n=88)	0 - 3	4 - 15	$\geq$ 16
12-16 (n=156)	0 - 3	4 - 11	$\geq$ 12

This particular sample includes about 340 children. This is somewhat lower than the number we have for the Delay Task scores because that Vigilance Task came along a bit later in the development process. While we have means and standard deviations available, we have presented percentiles because they are more readily understandable to practitioners without much experience with statistics. Also, the distribution tends to be skewed to the negative side (that is, children tend not to have many errors of commission) so the percentile is a more accurate reflection of the distribution. Our practice has been to divide the scores into ranges. These include ABNORMAL which is the 5th percentile or lower, BORDERLINE which is the 6th through 25th percentile, and NORMAL, which is 25th percentile and above.

Now, it would be nice to leave it all at this table and say that we've done what we need to do. But, of course, most of us do feel the need to go beyond the "arrogant" approach to examination of cross-validation for this measure.

First let me review briefly some of the scales used in these studies. For information from parents, we have come to rely on the Child Behavior Checklist (CBCL; Achenbach, 1978) because of its fine standardization and large pool of items and factors. In some of our standardization studies we have used the Conners Parent Questionnaire (Goyette, Conners & Ulrich, 1978).

As for Teachers, the Teacher Rating Form (TRF; Edelbrock & Achenbach, 1980) is the companion version of the Child Behavior Checklist. You will hear most about the Innattentive and Nervous-overactive Factors as these are most related to issues of attention deficit.

Another fine measure has been developed by Ullman, Sprague and Sleator (1984) at the University of Illinois. The ACTeRs scale generates 4 factors, including Attention, Hyperactivity, Social Skills and Oppositional behavior. We also have used the Teachers Rating Scale (Goyette, et. al., 1978), scored both by the original and IOWA (Loney & Milich, 1982) formats, in some of the data you will be viewing.

Before dealing with issues of cross-validation and the commission score let's take a look for a moment at the nature of the criteria we're using to judge agreement:

Intracorrelations Among TRF Factors

(n=60)

	Nervous-Overactive	Aggressive	Social Withdrawal	Anxious	Internalizing
Inattention	.83	.64	.61	. 28	.50
Nervous-Overactive		.58	.45	.20	.34
Aggressive			.60	-.09	.46
Social-Withdrawn				.61	.83
Anxious					.93

The first point to be drawn from our data and, as you know, from the work of Loney and others, has to do with the high degree of overlap among factors that supposedly represent relatively independent dimensions of behavior. For example, there is a high degree of relationship among ratings of attention, aggression, oppositional behavior and even anxious behavior on the Teacher Rating Form.

Let me point out for this table as well as for the series you are about to see that the group used was comprised of 354 nonhyperactive subjects. I am showing you these data

189

because it makes sense to see how parents and teachers in general rate children's behavior, before confounding the question with what happens when the children they are rating have been referred for psychiatric difficulties. Having said all that, the correlations among these scales, and the findings for the normal subjects very much hold when reanalyzed using clinical groups. While there is somewhat higher agreement in some areas, there is lower in others so the bottom line remains about the same.

You can see that the situation regarding overlap among factors is similar for the ACTeRS scale:

Intercorrelations Among Factors on the ACTeRS Scale

(n=198)

	Hyperactivity	SocialSkills	Oppositional
Attention	-.66	.73	.50
Hyperactivity		-.56	.63
Social Skills			-.56

Keep in mind that the Attention and Social Skills factors of the ACTeRS are scored in the opposite direction than the others. In other words, a high score on Attention indicates good attentional skills. Here are intercorrelations for Teacher's Conners:

Intercorrelations Among Standard & IOWA Conners' Factors

(n=291)

	I-O	Hyperactivity	Conduct Problem	Inattentive-Passive
IOWA				
Aggression	.76	.73	.65	.74
I - O		.93	.76	.77
Conners				
Hyperactivity			.68	.67
Conduct Problem				.43

1410

The following table illustrates the same point but examines the intercorrelations among the various teacher rating scales:

ACTeRS, TRS, IOWA & TRF Intercorrelations

	ACTeRS		
	ATTENTION	HYPERACTIVITY	OPPOSITIONAL
TRS (n = 145)			
Conduct Problem	-.55	.56	-.83
Inattentive-Passive	-.81	.59	.43
Tense-Anxious	-.37	.22	.25
Hyperactive	-.68	.79	.70
IOWA (n = 145)			
I - O	-.77	.75	.72
Aggression	-.62	.59	.58
TRF (n = 51)			
Inattentive	-.87	.84	.47
Nervous-Overactive	-.73	.79	.47
Aggressive	-.46	.82	.89
Anxious	-.13	.18	.08
Internalizing	-.31	.37	-.28
Externalizing	-.73	.92	.75

Again there is a high degree of overlap among scales of hyperactivity, aggression, oppositional behavior, and even anxiety.

This table also demonstrates, though, that there is a fair degree of agreement between the various factors from teacher scale to teacher scale. Remember that these are ratings by one teacher of one child's behavior using two scales designed to assess the same behavior.

The situation regarding overlap is no better for parent ratings (Table 1). In fact, they seem to be even more intertwined. Most striking is the high correlation between the Hyperactivity Factor and Social Withdrawal.

These kinds of findings have led some to conclude that there are few differences between conduct problems and hyperactivity. An alternative explanation is that raters tend not to be very fine-grained in their judgments of behavior and often will lump together behavioral domains which, in reality, are distinct. A third hypothesis is that the rating scales don't ask the right questions and, consequently, do not allow the rater sufficient opportunity for differentiation among categories of behavior.

191

One obvious conclusion to be drawn, though, is that criteria such as teacher and parent rating scales are, by definition, going to include in the sample large numbers of aggressive and/or oppositional and also anxious children. In other words, study groups constituted using rating scales, rather than being homogeneous, will most likely represent various subgroups of hyperactivity or of other disorders which can be interpreted by a teacher or parent as hyperactivity.

The impact on us as test developers is clear: it is problematic verifying the sensitivity and specificity of a test when the "accepted" criteria themselves tend to be relatively non-specific and will form diverse, grab-bag groups of hyperactive subjects.

In this next table you see what we also already know -- parents and teachers tend not to agree very closely on whether or not a child is hyperactive, or the degree to which they exhibit these kinds of behaviors:

Essential Intercorrelations Between Parent & Teachers on Measures of Hyperactivity/Attention

TEACHERS		PARENTS CBCL – Hyperactivity
TRF	Inattentive	.47 (60)
	Nervous-Overactive	.40 (60)
TRS	Hyperactivity	.35 (265)
IOWA	I - O	.39 (264)
ACTeRS	Attention	-.43 (172)
	Hyperactivity	.31 (172)

Let me remind you that the TRF and CBCL scales contain almost identical items. The correlations between the CBCL Hyperactivity factor and the other scales related to hyperactivity also fall in a very modest range, particuarly when you consider that parent and teacher are rating the same child with instruments designed to assess the same domain of behavior. The situation appears worse for the two Conners scales. The correlations between the Parent and Teacher Hyperactivity Index was .25 (n=49).

Agreement between parent and teachers is particularly low when you ask them to rate aggressive and oppositional

192

behavior:

Correlations Between Parent & Teachers on
Measures of Aggression/Opposition
TEACHERS

	TRS Conduct Problem	IOWA Aggression	ACTeRS Oppositional	TRF Aggression
PARENTS CBCL				
Aggression	.22 (265)	.22 (264)	.29 (173)	.24 (60)
Cruel	.20 (131)	.26 (131)	.29 (173)	.03 (25)

What is even more interesting is how the level of agreement between parents (CBCL) and teachers (ACTeRS) changes depending on the rating scale and the age of the child:

Intercorrelations Between Parent and
Teacher Ratings by Age

(AGE IN YEARS)	CBCL HYPER. & ACTeRS HYPER.	CBCL HYPER. & TRF INATT.	CBCL AGGR. & TRF C-P
6 - 7	.46 (36)	.37 (52)	.42 (52)
8 - 9	.47 (61)	.29 (87)	.19 (87)
10 - 11	.12 (42)	.39 (65)	.22 (65)

The correlation between the CBCL Parent Hyperactivity Factor and the Acter's Hyperactivity Factor, for example, ranges from .12 to .46 -- depending on the age of the child.

My intent in reviewing these data on the rating scales is not to deride them. Those of you familiar with our studies know that some of my best empirical friends are rating scales. I did want to underline a point made so often in the literature but sometimes forgotten. That is, that rating scales are far from being definitve criteria for constituting clinical groups. As you will see, you can select vastly different groups depending on what factors you employ and which raters or combination of raters you choose to rely upon. It is because of this state of affairs that I suppose I've come to bristle a bit at the question, "How accurate is the GDS" or the statement, "The GDS misses x percentage of children". Those statements imply that there is a definable target against which one could judge accuracy

193

and that the pool of hyperactive children is a relatively homogenous one. It is also why, as a test developer, you constantly deal with being damned if you do and damned if you don't. When the GDS scores correlate with rating scales I will hear "That's not good because the rating scales are so inaccurate." When the correlations are low I hear, "Well, the GDS can't be very accurate because they don't correlate with the rating scales." The statisticians, in a way, resolve the entire issue by telling me that correlations are meaningless altogether.

With renewed respect for the issue of criteria confusion, we now move on to data regarding the validation of the Commission score. Presented in Figure 1 are median Total Commission scores for three groups: Children rated hyperactive by parents and teachers, children classified as non-hyperactive emotionally disturbed (ED), and nonhyperactive learning disabled (LD). The data are from dissertation studies by Peter Oppenheimer of University of Virginia and Deborah DiNiro of Syracuse University. Both the ED and LD groups are defined by placement. The LD children were all placed in special programs because they met New York State criteria for a learning disability. Almost all these children were reading-disabled. The ED group were all in psychiatric treatment at least once weekly. The bulk of the group were severely disturbed children who attended a day treatment program. You should keep in mind the level of severity of psychopathology in considering these data.

You will see in this figure that the Commission score differentiates clearly between these three major groups. Compared to the normal group, there is a dramatic difference in the median score -- from 13.5 for the hyperactive group to 4 for normals. These differences between the hyperactive group and the other two categories are very significant when analyzed statistically.

While things are looking good for the Commission score, it is only fair to present this next figure (see Figure 2). These are the various medians for hyperactive groups comprised according to a variety of criteria. That is, consensus hyperactives upon whom there is agreement between parent and teacher, hyperactives classified by parent or teacher, and those selected when you rely only on the parent or only on the teacher.

Regardless of how you comprise the groups, there is a difference between hyperactive and nonhyperactive. Nonetheless, you can also see that there are some very significant differences in median scores depending on what criteria you employ. Just look at the N for each group when you pick out "consensus hyperactives" as opposed to parent or teacher, or parent alone or teacher alone rated hyperactives.

We run into the same ambiguity when you ask the next logical question -- How many of the hyperactive group will be identified by the GDS as hyperactive?

Identification by Total Commission Score of Hyperactive Groups Constituted by Various Criteria

Rated ADD-H by:

	PARENT & TEACHER	PARENT OR TEACHER	PARENT ONLY	TEACHER ONLY	TEACHER ACTeRS
(AGE IN YEARS)					
6 - 7	71%	58%	62%	64%	--
8 - 9	42%	41%	46%	42%	-
10 - 11	72%	47%	45%	68%	--
TOTAL	61%	48%	51%	58%	68%
	(33/54)	(59/122)	(54/106)	(39/68)	(13/19)

Here you see rates of agreement when you select out from variously-constituted samples children who fall at the 10th percentile or lower on the different scales. (This level was used to allow for a sufficient N). We also used the 10th percentile for the GDS Commission score. You will note that the levels of agreement vary depending on the criteria for classification, the rating scale used, and the age of the child. Depending on your point of view, the GDS "misses" around 30% of hyperactive children or the parents and/or teachers overrated 30% of the sample as hyperactive.

To put this issue in perspective, I have here the same kind of exercise but this time we will look at the percentage of children picked out by the Acters Hyperactivity and Attention factors when applied to variously constituted groups of ADD/Hyperactive children:

Identification by ACTeRS Scales of Hyperactive Groups Constituted by Various Criteria

RATED ADD-H BY:

	PARENT AND TEACHER	PARENT OR TEACHER
<u>ACTeRS</u>		
Hyperactive	30%	18%
	(16/54)	(22/122)
Attention	54%	56%
	(29/54)	(68/122)

Remember the high correlations between the ACTeRS and the Achenbach scales and the fact that the Acters and TRF are filled out by the same teacher. You can see here that the levels of agreement are very low. In fact, they are generally lower than those achieved by the GDS.

195

The next concern regards the issue of false positives. How many nonhyperactive children will the commission score identify?

Identification of Nonhyperactives by Total Commissions and ACTeRS

	SCHOOL NORMALS	COMBINED SCHOOL NORMALS & CLINIC NON-ADD
	(n = 50)	(n = 89)
TOTAL COMMISSIONS	12%	20%
ACTeRS		
Hyperactive	2%	6%
Attention	28%	34%

You will note that, at the 10% level, it selected 12% of a separate non hyperactive sample. You can see how that compares to the Acters scales. The clinic nonhyperactive sample includes the normal subjects plus those children rated by both parent and teacher as nonhyperactive. Out of 211 subjects, only 39 were rated by both as nonhyperactive. Incidentally, approximately 24% of the parents of 114 normal children in one of our studies responded that their children were overactive or impulsive.

So, at this point, despite all the caveats and complexities, you can see that this one score, Total Commissions, holds up reasonably well.

Now we move on to the correlations between Total Commissions and other lab measures:

Correlations Among Total Commissions, Fruit Distraction Test, VADS & Other Child-Based Measures

	TOTAL COMMISSIONS	VADS TOTAL	FRUIT 4 - 1
FDT 4 - 1	.34	.41	
VADS TOTAL	.40		
TOTAL ER	.31	.10	.18
VTWD-CORR.	-.37	.57	-.30
VTWD-COMM.	.56	-.40	.33
VTAUD-CORR.	.18	.48	-.43
WRAT %	-.11	.41	-.26

The "FTD" stands for Fruit Distraction Test and was developed by Santostefano (1971). It is a kind of Stroop Color Distraction Test for children and requires them to name colors surrounded by various degrees of distractors and contradictions.

The Visual Aural Digit Span Test (VADS) was developed by Koppitz (1977) as a measure of intersensory integration and

196

memory. The child is required to repeat digits presented either visually or aurally. The child's response is given either verbally or in written form. The Total Score represent an overall index of performance.

The next measures are correct and commission scores for two GDS enhancements we are developing. The Vigilance Task with Distracters (VTWD) is identical to the regular Vigilance Task except that random digits flash at random intervals on the outer two positions on the LED diplay. The Auditory Vigilance Task (VTAUD) is also identical to the standard Viglance Task except the digits are presented orally by means of a speech synthesizer.

Finally, the Wide Range Achievement Test (or WRAT) (Jastak & Jastak, 1978) generates an estimate of academic achievement.

The table indicates a modest degree of agreement among the various measures. Most of the correlations are statistically significant.

Conclusions:

Let me emphasize at the outset of my concluding remarks that my intent has not been to suggest that the GDS scores are better than rating scales or should supplant clinical judgment. Nor do I mean to imply that the situation is so arbitrary that despair and cynicism should abound. My main point is that, with such an important diagnosis, where there reigns so little consensus about diagnostic practice, and where there are criteria of such tenuous validity, it makes sense to throw out as wide a psychometric net as possible.

The inclusion of a laboratory measure, to me, appears essential to a sophisticated evaluation for ADD/Hyperactivity. If nothing else, it gives the clinician an opportunity to see for him or herself how a child performs in a standard situation which requires delay and sustained attention. Without behavior-based data, we are far more likely to overdiagnose this disorder, especially with boys. My hunch is that, more often than not, current practices tend to overlook girls with significant deficits in attention in the context of more manageable physical behavior

I support use of the GDS tasks because they are reliable, easy to use, well-standardized and well-supported by research efforts. Indeed, the feedback we get from practitioners is very positive and the validational studies are certainly encouraging.

As I mentioned at the outset, though, my point goes beyond issues regarding the GDS, per se.

These data re-emphasize what has been salient in the literature. The three major sources of data, that is information from parents, from teachers, and from the child him or herself via the lab measures tend to overlap only marginally. The child who emerges as abnormal in all three domains is going to be a rare, and probably quite difficult bird. While this consensus hyperactive has everybody agreeing, there is, of course, the questions of whether those other children who are considered hyperactive by two of the three realms could not benefit from treatment and might not represent important subgroups of the ADD/Hyperactive population.

These sort of data point to the need for us to become more intent on identifying subgroups of ADD/Hyperactive children beyond the "with and without hyperactivity" distinction. Delineations by "with and without aggression" make a great deal of sense but I am not sure we have good technology to make that distinction reliably. It was discouraging to us that the two scales of the IOWA Conners, that is the one presumably more related to inattention and the one related more to aggression, were so highly correlated.

I suppose one could become thoroughly cynical and suggest that the diagnostic categories should be ADD as rated by parent, ADD as rated by parent and teacher, etc.

These kinds of findings have convinced us all the more of the utility of the criteria offered by Dr. Barkley because they do throw out that wide net in an effort to establish severity, chronicity, and pervasiveness:

Diagnostic Criteria for ADD/Hyperactivity
(Adapted from Barkley, 1981)

1. Parent or teacher complaints of poor attention span, impulsivity, restlessness and inability to restrict behavior asa situation or adult demands.

2. Complaints of behavior place child 2 standard deviation (i.e. in 5th percentile from mean for his/her age and sex group as determined by well-standardized behavior scale of parent or teacher opinion.

3. BORDERLINE/ABNORMAL scores on either GDS task.

4. Onset prior to 5 years, 11 months (differs from DSM-III).

5. Duration of symptoms at least 1 year.

6. Pervasiveness of symptoms such that the child is scored at or below the 5th percentile for either the Home Situation Questionnaire (HSQ, 5th = 10) or School Situation Questionnaire (SSQ, 5th = 9).

7. IQ > 70.

8. Child does not display symptoms of autism or psychosis, or show evidence of blindness, deafness, aphasia or gross neurological disease (e.g. tumors, strokes, neurodegenerative disease, or obvious CNS trauma).

As you can see from this table, we would add BORDERLINE or ABNORMAL scores on the GDS tasks as additional criteria. We also have begun using different data for the Home Situations Questionnare and School Situations Questionnaire. Our data on 85 normal subjects suggest that the cuttoff for each should be 10 for the HSQ and 9 for the SSQ.

Finally, anyone concerned about the place of clinical

198

judgment in this process should take great heart from the current situation. There is still plenty of room for sound judgment related to the integration of data and the appropriateness of conclusions. The complexity inherent in child behavior problems will, I am sure, guarantee the value of clinical skills for a long time to come.

References

Achenbach, T.M. (1978) The child behavior profile: I. Boys aged 6-11. Journal of Consulting and Clinical Psychology, 46, 478-488.

Barkley, R.A. (1981) Hyperactive children: A handbook for diagnosis and treatment . New York: The Guildford Press.

Edelbrock, C. & Achenbach, T.M. (1984). The teacher version of the Child Behavior Profile I: Boys aged 6 - 11. Journal of Consulting and Clinical Psychology , 52,(2), 20744-217.

Gordon, M. (1982). The Gordon diagnostic system, Littleton, Colorado: Clinical Diagnostics, Inc.

Goyette, C.H., Conners, C.K. & Ulrich, R.F. (1978). Norm ative data on revised Conners parent and teacher rating scales. Journal of Abnormal Child Psychology, 6, 221-236.

Jastak, J.F. & Jastak, S. (1978). Wide range achievement test: 1978 revised edition. Wilmington: Jastak Associates.

Koppitz, E.M. (1977). The Visual Aural Digit Span Test. New York: Grune & Stratton.

Loney, J. & Milich, R. (1982). Hyperactivity, inattention, and aggression in clinical practice. Advances in Developmental and Behavioral Pediatrics, 3, 113-147.

Rosvold, H., Mirsky, A., Sarason, I., Bransone, E. & Beck, A. (1956). A continuous performance test of brain damage. Journal of Consulting Psychology, 20, 343-352.

Santostefano, S. (1971). Fruit distraction test: A procedure for assessing in children the cognitive principle of "field articulation". Boston University School of Medicine, Boston, Mass: Author.

Ullmann, R.K., Sleator, E.K. & Sprague, R.L. (1984). A new rating scale for diagnosis and monitoring of ADD children. Psychopharmacology Bulletin, 20,(1), 160-164.

TABLE I

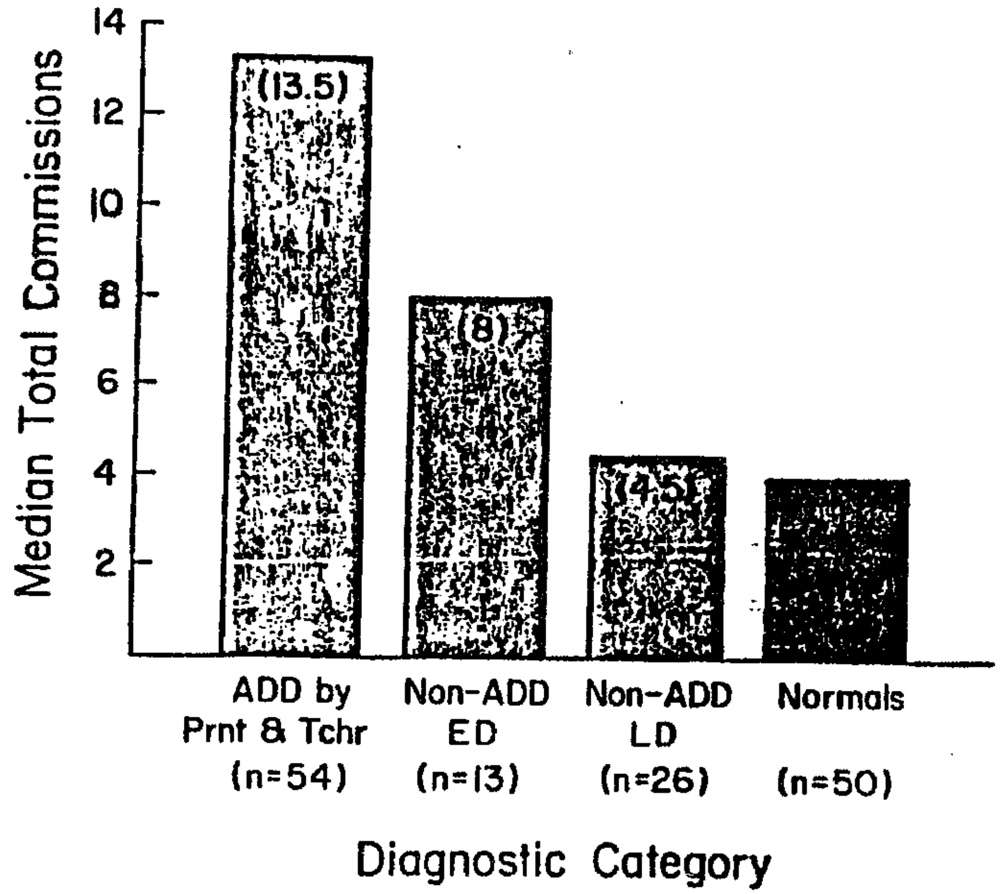
INTRACORRELATIONS AMONG CBCL FACTORS

(N = 345)

	AGGRESSION	HYPERACTIVE	CRUEL	ANXIOUS	SOCIAL WITH.	INTERNAL	EXTERNAL
GRESSION		.75	.72	.42	.60	.77	.97
PERACTIVE			.56	.44	.60	.75	.86
RUEL				.40	.43	.52	.72
IXIOUS					.41	.65	.45
ICIAL WITHDRAWAL						.76	.64
ITERNAL							.81

FIGURE I

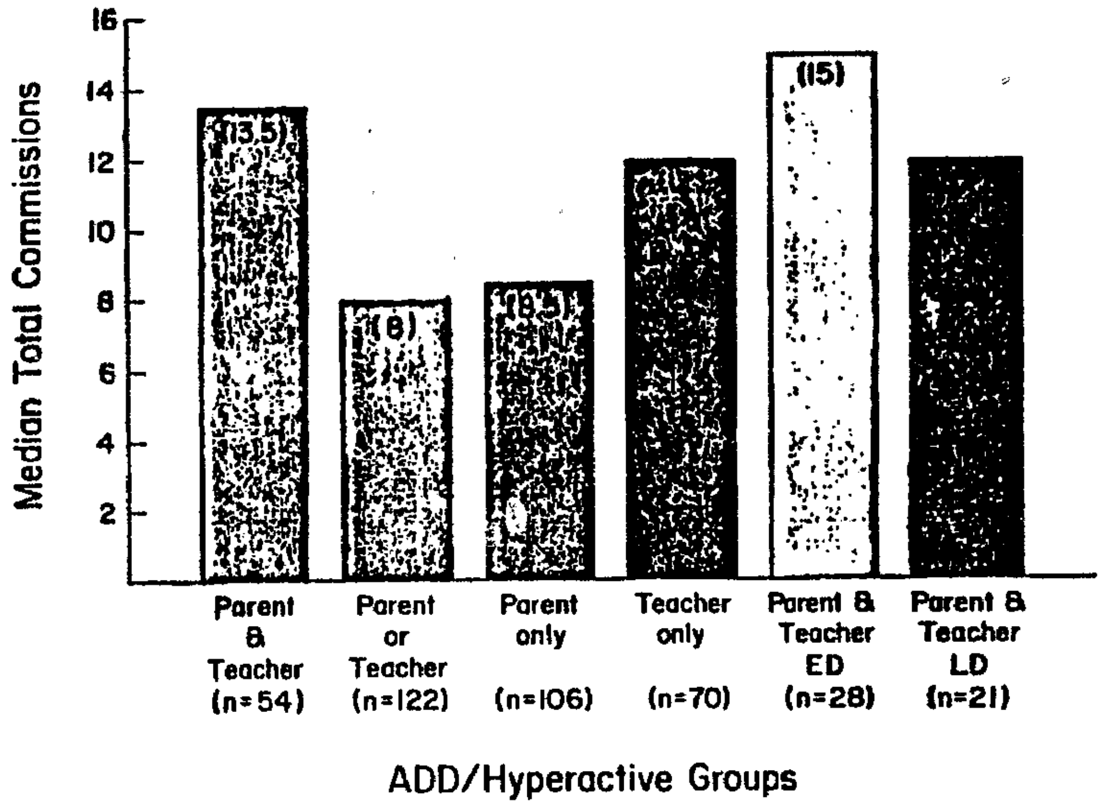
Median Total Commission Score by Diagnostic Category



202

FIGURE II

Median Total Commission Scores for ADD/Hyperactive Groups Formed by Various Criteria



203

RS232C communications port. Software is available which allows for the direct transmission of GDS data to a microcomputer where it can be tabulated, graphed, stored, and compared to normative data. Connection of the GDS to a microcomputer enables collection of some ancillary data which cannot be extracted from the stand-alone unit.

The portability and ease-of-operation of the GDS has allowed for large-scale data collection. The standardization sample is comprised of 900 boys and girls from 3 to 16 years of age (Gordon & Mettelman, 1985). An additional 1,100 hyperactive and nonhyperactive protocols from various subject populations, including the deaf, blind, Spanish-speaking, emotionally disturbed, and learning disabled, have also been gathered.

A series of validational studies has shown that these game-like tasks differentiated accurately between hyperactive and nonhyperactive children from both outpatient, and day treatment settings (Gordon, 1979; McClure & Gordon, 1984). In a sample of school-referred children, the GDS distinguished children with ADD from those classified as reading disabled, overanxious, and normal (Gordon & McClure, 1983). Over 20 university and medical center research sites are currently conducting investigations involving the GDS. Most studies concern the effectiveness of the GDS for determining drug responsivity and for evaluating the success of pharmacotherapy (Atkinson, 1985; Barkley, 1985; Rapport, 1985). Others are investigating the use of the GDS for evaluating children with known brain damage, preschoolers considered "at-risk" for impulsive behavior, and the relationship between GDS scores and observational measures of classroom behavior. The GDS is also being used by school systems and pediatric practices across the country for the clinical evaluation of ADD-H.

Comments and Conclusions

A chronic impediment in the field of ADD-H has been the paucity of universal criteria and procedures for subject selection and treatment monitoring. Even in the realm of behavioral assessment, where truly creative approaches to testing children have been developed, procedures have generally not been sufficiently practical or well-standardized to allow for acceptance by the majority of professionals (Loney, 1981). As such, most continue to rely on rating scales or tests, such as the Wechsler Intelligence Scale for Children (WISC-R) (Wechsler, 1974), which have been shown to be of limited usefulness in differen- [illegible] ance of reliable administration and valid interpretation. The growing acceptance of the GDS within the professional community also permits greater comparison of data across studies.

It must be emphasized, however, that the GDS was never intended as a divining rod for ADD-H. This is a complex disorder which represents an array of subgroups and interactive diagnostic dimensions. The GDS is seen as an important tool to be used only in conjuction with other selection criteria and clinical judgment.

References

Atkinson, A.W., Cohen, P.C., and Kelly, P.C. Attention deficit disorder: The effects of ritalin on self-esteem a comparison of ACTeRS teacher scale, Conners' parent scale, and Gordon Diagnostic System in diagnosis and management. Paper presented at the American Academy of Pediatrics Meeting, Atlanta, Georgia, April 1985.

Barkley, R.A. Assessment of stimulant drug responding in ADD-H children. Paper presented at the 93rd Annual Convention of the American Psychological Association, Los Angeles, California, August 1985.

Conners, C.K. *Continuous Performance Test* [Computer program]. Kensington, Maryland: Behavioral Medicine, Inc., 1980.

Davenport, W. Vigilance and arousal: Effects of different types of background stimulation. *J. Psychol. 82*, 339–346, 1972.

Douglas, V.I. Stop, look and listen: The problem of sustained attention and impulse control in hyperactive and normal children. *Can. J. Behav. Sci.*, 4:259–282, 1972.

Doyle, R.B., Anderson, R.P., and Halcomb, C.G. Attention deficits and the effects of visual distraction. *J. Learn. Disab.*, 9:59–65, 1976.

Gordon, M. The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive children. *J. Abnorm. Psychol.*, 7:317–326, 1979.

Gordon, M., and McClure, F.D. The objective assessment of attention deficit disorder. Paper presented at the 91st Annual Convention of the American Psychological Association, Anaheim, California, 1983.

Gordon, M., and McClure, F.D. Assessment of attention deficit disorders using the Gordon Diagnostic System. Paper presented at the 92nd Annual Convention of the American Psychological Association, Toronto, Canada, 1984.

Gordon, M., and Mettelman, B.B. *Threshold Tables for the Gordon Diagnostic System.* (Available from Clinical Diagnostics, Inc., 300 E. Mineral Avenue, Suite 6, Littleton, Colorado 80122); 1985.

Hiscock, M., Kinsbourne, M., Caplan, B., and Swanson, J.M. Auditory attention in hyperactive children: Effect of stimulant medication on dichotic listening performance. *J. Abnorm. Psychol.*, 88:27–32, 1979.

Loney, J. Evaluating treatments for childhood hyperactivity: Some methodological considerations. In K.D. Gadow, K.D., and S. Loney, J. (eds.), *Psychosocial Aspects of Drug Treatment for Hyperactivity.* Boulder, Colo.: Westview Press, 1981, pp. 77–103.

Klee, S.H., and Garfinkel, B.D. The computerized continuous performance task: A new measure of [illegible]

allow for acceptance by the majority of professionals (Loney, 1981). As such, most continue to rely on rating scales or tests, such as the Wechsler Intelligence Scale for Children (WISC-R) (Wechsler, 1974), which have been shown to be of limited usefulness in differentiating between groups (Douglas, 1972). Furthermore, comparison of studies across research programs has often been hampered by idiosyncratic measures or sets of criteria.

The GDS represents an effort to establish a standard procedure for evaluating certain aspects of attention and self-control. While it is more costly and, in certain respects, less flexible than software-driven approaches, its portability, ruggedness, and extensive base of normative data offer the user greater assur-

Loney, J. Evaluating treatments for childhood hyperactivity: Some methodological considerations. In K.D. Gadow, K.D., and S. Loney, J. (eds.), *Psychosocial Aspects of Drug Treatment for Hyperactivity*. Boulder, Colo.: Westview Press, 1981, pp. 77–103.

Klee, S.H., and Garfinkel, B.D. The computerized continuous performance task: A new measure of inattention. *J. Abnorm. Child Psychol.*, 11(4):489–496, 1983.

Margolis, J.S. *Academic Correlates of Sustained Attention.* (Unpublished thesis, University of California, Los Angeles, 1972).

McClure, F.D., and Gordon, M. The performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. *J. Abnorm. Child Psychol.*, 12(4):561–572, 1984.

Rapport, M.D. Comparing classroom and clinic measures of ADD: Dose response effects. Paper presented at the 93rd Annual Convention of the American Psychological Association, Los Angeles, California, August 1985.

Rosvold, H.E., Mirsky, A.F., Sarason, I., A continuous performance of brain damage. *J. Consult. Psychol.*, 20:343–350, 1956.

Weschsler, D. *Weschsler Intelligence Scale for Children-Revised.* New York: Psychological Corporation, 1974

Microprocessor-Based Assessment of Attention Deficit Disorders

Michael Gordon, Ph.D.[1]

Dissatisfaction with formulating a diagnosis of ADD with hyperactivity (ADD-H) based almost entirely upon clinical judgment or the perception of raters has spawned the development of behavior-based assessment procedures. These efforts typically involve administration of the Continuous Performance Test (Rosvold, et al., 1956), or related measures of attention and self-control (Davenport, 1972; Doyle et al., 1976; Hiscock et al., 1979; Margolis, 1972). Although bulky and expensive electromechanical devices had previously been required, researchers are now programming microcomputers to administer these laboratory tasks to children (Conners, 1980; Klee & Garfinkel, 1983).

Use of the microcomputer has certain advantages, including flexibility of administration, the ability to store multiple data points, and, if a microcomputer is already available, cost-effectiveness. This approach, however, does have drawbacks for the researcher and, in particular, the clinician. The size of even portable microcomputers can make testing in multiple sites cumbersome. It is perhaps for this reason that standardization samples for software-driven programs tend to be, at best, limited in number and breadth. The transport of disk drives, monitors, keyboards, and the computer, itself, from location to location often discourages use of the procedure in other than a single research or clinic setting.

Another potential limitation of microcomputer-based testing stems from concerns surrounding the reliability of administration. Computer monitors vary in the size, shape, intensity, and color of displayed characters. Each brand of computer presents to the child a different keyboard, and accepts different types of peripherals, such as joysticks, which are often used as manipulanda. Without con- [illegible] factors as the amount [illegible] ministration of the task. Efforts to secure the microcomputer are often unsatisfactory because they usually involve additional hardware, cabling, and expense.

An alternative approach to testing attention and self-control has been developed by the author (Gordon & McClure, 1983, 1984). The goal of this project has been to establish a practical, reliable, and well-standardized procedure available to both researchers and clinicians. The Gordon Diagnostic System (GDS) is a microprocessor-based, portable unit which, without an external microcomputer, allows for the administration of multiple tasks. The Delay Task, based upon a DRL operant schedule requires the child to inhibit responding in order to gain a reward. Specifically, the child is instructed to press a button, wait, and then press the button again. If she or he refrains from responding for at least 6 seconds, a light flashes and a reward counter increments. If the child responds before the interval elapses then the timer resets and no reward points are recorded. The GDS Delay Task generates three major scores: the number of responses (button presses), the number of correct responses (i.e., Correct) and the Efficiency Ratio which represents the percentage of correct responses.

The GDS also contains the Vigilance Task, which is a version of the Continuous Performance Test (Rosvold et al., 1956). The child is presented with a series of digits on the front display and is told to press the button every time a "9" appears that had been preceded by a "1". The GDS records the number of correct presses, the number of times the child fails to press the button when the "1/9" combination appears (i.e., errors of omission), and the number of extraneous button presses (i.e., errors of commission). For the testing of younger children, the GDS contains a "1" mode, which requires the subject to press the button every time a "1" appears. The same performance measures are recorded.

Although normative data were gathered using standard settings for task parameters, the design allows the user to select a wide range of parameters. This feature enabled the modification of parameters for the testing of adults as well as very young children.

Because the GDS is controlled by a programmable microprocessor, additional tasks can be explored. One such enhancement being studied is a version of the Vigilance Task which assesses the impact of distraction on a child's ability to sustain attention. This Distractibility [illegible]

...ability of administration. Computer monitors vary in the size, shape, intensity, and color of displayed characters. Each brand of computer presents to the child a different keyboard, and accepts different types of peripherals, such as joysticks, which are often used as manipulanda. Without control over such critical factors as the amount of pressure necessary to actuate a switch, or the accuracy of timed sequences, repeatable administration across multiple microcomputers cannot be assured.

Finally, ADD-H children, despite a reputation for academic underachievement, routinely display a fine facility for disassembling delicate equipment. Unless the examiner is immediately present, the highly impulsive child will inevitably stick a finger in disk drives, unplug cables, adjust monitors, or in some other fashion interrupt standard ad-

...parameters for the testing of adults as well as very young children.

Because the GDS is controlled by a programmable microprocessor, additional tasks can be explored. One such enhancement being studied is a version of the Vigilance Task which assesses the impact of distraction on a child's ability to sustain attention. This Distractibility Task is identical to the standard Vigilance Task except that random digits flash at random intervals on the outer positions of the LED display. The subject is still required to press the blue button when a "9" comes right after a "1". The only difference is that numbers flash on either side of the center (i.e., relevant) digit. Other enhancements under development are designed to evaluate attention across sensory modalities.

Even though the GDS administers tasks independently of a microcomputer, it can communicate with external hardware via an

[1]State University of New York, Syracuse, NY 13210.

*Pharmacology Bulletin (in press) Vol. 22 (1).

CHAPTER IN J.LONEY (ED.).
THE YOUNG HYPERACTIVE CHILD
HAWORTH PRESS (IN PRESS)

HOW IS A COMPUTERIZED ATTENTION TEST USED IN THE DIAGNOSIS OF ATTENTION DEFICIT DISORDER?

Michael Gordon, Ph.D.

SUNY/Upstate Medical Center

The advent of computerized assessment of behaviors associated with Attention Deficit Disorders (ADD) represents another step forward in the effort to define reliable and meaningful diagnostic criteria. These techniques were born of concern about the extensive degree to which diagnostic decisions were founded upon opinion, or upon data from traditional psychological tests of limited relevance to issues surrounding this disorder. Although clinical judgment, behavior rating scales, and clinical interviews are critical to a sophisticated evaluation, each approach harbors well-documented limitations. In the face of considerable evidence to the contrary, one would be hard put to justify the statement "I know an ADD child when I see one" or the practice of using only rating scale scores or IQ factors to formulate a diagnosis.

While not intended as magic geiger counters for ADD, computer-based laboratory measures do offer both the researcher and clinican an opportunity to incorporate data derived from a child's actual behavior. Unlike the other clinical methods, laboratory tasks generate objective data about a child's ability to perform in situations tailored to strike at the characteristic weaknesses of an ADD child. As such, they generally require a child to sustain attention and control behavior over a period of time and with varying

208

degrees of feedback regarding performance. The tasks are often variants of attentional measures, such as the Continuous Performance Test (Rosvold, Mirsky, Sarason, Bransome & Beck, 1956), in which a child must respond only to a specific combination of symbols in a stream of irrelevant symbols. Many years of research indicate that ADD children fare poorly when required to attend and to develop a strategy for self-control.

Laboratory measures have had a long history in ADD research and a very short one in actual clinical practice. Research versions have traditionally been bulky, expensive, and impractical. There has also existed a certain degree of resistance to using mechanical measures within an evaluation which traditionally has relied heavily on the clinician's judgment. The burgeoning growth of computer technology has solved the technological limitations of earlier electromechanical devices. The increasing popularity of computers in clinical practice has also led to greater comfort with computerized assessment of behavior.

Considerations involved in selecting a computerized assessment technique are essentially identical to those important for the evaluation of any psychological test. The computerization of a measure does not obviate the need for evidence concerning reliability of administration, test-retest reliability, robust standardization, or meaningful studies of validity. Most of the available techniques have enjoyed only limited research investigation of psychometric properties. These procedures have generally been standardized on small groups of subjects, usually children who have been referred to a psychiatric clinic but for reasons other than those associated with ADD. While such samples may suffice for certain research purposes, the clinician is interested more in how a particular child's performance compares to that of peers than to the scores of children referred for a range of psychiatric problems.

Extensive standardization of computerized assessment procedures are particularly critical in the light of the legendary variability inherent in the test performances of children. It is also important to keep in mind that most of the scores generated by these measures are age-related. Without a very substantial normative base, misdiagnosis can result from failure to adjust for a child's developmental status.

Another key consideration concerns the practicality of an assessment procedure. If a technique is not designed to fit comfortably into the daily practice of a busy clinican, it will likely fall from use regardless of its diagnostic efficacy. While software-driven assessment programs have been available for many years, they generally have not found their way into clinical practice, often because clinicians have considered them too cumbersome to administer and score. Most practitioners have neither the time nor the computer expertise to struggle with complicated instructions or lengthy procedures. Most of the techniques also fail to meet the clinican's need for portability. Since the majority of those involved in serving children referred for ADD travel among schools and/or offices, they are loathe to carry around bulky equipment or to expend effort in connecting cables and attaching peripherals.

The clinician would, therefore, be wise to select a procedure that can be employed with ease and confidence. Along these lines, the assessment program should be accompanied by clear operating instructions, support materials regarding interpretation, technical data, and active service and research support.

Unfortunately, there are only a handful of available computerized techniques which even begin to approach the criteria mentioned above. Most rely on microcomputers to generate the tasks (Conners, 1980; Klee & Garfinkel, 1983; Greenberg, 1985) in contrast to a self-contained, microprocessor-based unit dedicated to task administration (Gordon, 1982). The advantage of the former strategy is that these software programs tend to be less expensive, more flexible in the varieties of tasks that can be administered, and more amenable to data storage. However, they also tend to be less transportable, rugged, practical, and reliable in administration. As indicated above, software packages tend to be poorly normed if normed at all. The one program that has enjoyed some standardization is the Garfinkel Assessment Battery, but this technique is restricted to research purposes.

The most widely-used procedure is the Gordon Diagnostic System (GDS; Gordon, 1982), developed by the author. Although supported extensively by research efforts, the GDS was designed specifically for clinical use. The GDS is a microprocessor-based, portable unit which, without an external microcomputer, allows for the administration of multiple tasks. The Vigilance Task requires the child to inhibit responding under conditions that make demands for sustained attention. A series of digits flashes one at a time on an electronic display. The child is told to press the button every time a "1" is followed by a "9". The GDS records the number of correct reponses, the number of time the child failed to respond to the "1/9" combination, (i.e. Errors of Omission), and the number of extraneous button presses (i.e. Errors of Commission). For the testing of younger children, the GDS contains a "1" mode, which requires the subject to press the button every time a "1" appears. The same performance measures are recorded.

The Delay Task requires the child to inhibit responding in order to earn points. Specifically, the child is instructed to press a button, wait a while, and then press the button again. If s/he refrains from responding for at least 6 seconds, a light flashes and a reward counter increments. If the child responds before the interval lapses then the timer resets and no reward points are recorded. The Delay Task yields three primary scores: the number of responses (button presses), the number of correct responses (i.e. Correct) and the Efficiency Ratio, which represents the percentage of correct responses.

The administration of both tasks takes less than twenty minutes. Although normative data were gathered using standard task parameters, the design allows the practitioner to select a wide range of settings. This feature enabled the modification of parameters for the testing of adults as well as very young children. The internal microprocessor generates the tasks and records quantitative features of a child's performance both for the entire session as well as for individual time blocks. In this way, the pattern of a child's performance across the session can be analyzed.

Because the GDS is controlled by a programmable microprocessor, additional tasks can be explored. One such enhancement being studied is a version of the Vigilance Task which assesses the impact of distraction on a child's ability to sustain attention. This Distractibility Task is identical to the standard Vigilance Task except that random digits flash at random intervals on the outer positions of the electronic display. The subject is still required to press the blue button when a "9" comes right after a "1". The only difference is that numbers flash on either side of the center (i.e. relevant) digit. Other enhancements under development are designed to evaluate attention across sensory modalities.

The portability and ease-of-operation of the GDS has allowed for large-scale data collection. The standardization sample is comprised of 1200 boys and girls from 3 to 16 years of age (Gordon & Mettelman, 1985). An additional 1100 hyperactive and nonhyperactive protocols from various subject populations, including the deaf, blind, emotionally disturbed, learning disabled, and Spanish-speaking have also been gathered. A series of validational studies has shown that these game-like tasks differentiated accurately between hyperactive and nonhyperactive children from both outpatient and day treatment settings (Gordon, 1979; McClure & Gordon, 1984). In a sample of school-referred children, the GDS distinguished children with ADD from those classified as reading disabled, overanxious, and normal (Gordon & McClure, 1983). Over 20 university and medical center research sites are currently conducting investigations involving the GDS. Most studies concern the effectiveness of the GDS for determining drug responsivity and for evaluating the success of pharmacotherapy (Atkinson, 1985; Barkley, 1985; Rapport, 1985). Others are investigating the use of the GDS for evaluating children with known brain damage, preschoolers considered "at-risk" for impulsive behavior, and the relationship between GDS scores and observational measures of classroom behavior. The GDS is also being used by school systems, mental health professionals, and pediatricians across the country for the clinical evaluation of ADD/Hyperactivity and assessment of therapeutic outcome.

Feedback from clinicians indicates that GDS testing tends to be conducted early in the diagnostic phase in an effort to achieve a more efficient evaluative process. The rationale for this approach is that GDS results can help direct the course of the ensuing evaluation by initially screening for general levels of self-control and attentiveness. For those children who perform within normal limits on the GDS and brief rating scale measures of hyperactivity, further assessment can be geared toward examining other possible explanations for the child's behavior aside from ADD/Hyperactivity, per se. Conversely, the clinician is likely to pursue more intently consideration of a diagnosis of ADD/Hyperactivity for those children whose GDS scores fall in the abnormal ranges.

Practitioners also report that the GDS testing to be most useful in helping to rule out a diagnosis of ADD/Hyperactivity for the many children who are not hyperactive and would best benefit from treatments other than those applied for ADD/Hyperactivity. The procedure is also used extensively in monitoring the effectiveness of pharmacotherapy. Finally, clinicians have consistently reported that, quantitative scores aside, the opportunity to observe a child perform in situations demanding of attention and self-control has been valuable.

Clinical examples:

Presented below are a series of cases in which GDS testing was included within an evaluation (adapted from Gordon, 1984, 1985). While a few case histories, truncated in their description because of space limitations, cannot convey the full impact of objective testing, they can serve to illustrate the sorts of instances where computerized assessment can be meaningful.

Case 1

A nine year old boy was referred with complaints of restlessness, noncompliance, fighting with other children, and poor academic performance. While his parents had been recently divorced, these behaviors were longstanding and had been problematic since early childhood. A medical examination by his pediatrician was unremarkable, and he was referred for a psychological evaluation.

Upon initial contact with the psychologist, this youngster appeared inhibited, quiet, and withdrawn, giving no indications of impulsive or hyperactive behavior. However, when confronted with the demands of the Delay Task, the boy was unable to maintain the facade. He achieved an Efficiency Ratio of .44, which falls in the Abnormal Range. His behavior throughout the Delay Task was disjointed, involving much out-of-seat activity and extreme restlessness. While these behaviors in some ways helped him to suppress responding, because he was otherwise occupied, he was nonetheless unable to refrain from emitting a large number of unreinforced responses. This pattern of behavior was repeated during the Vigilance Task. He consistently responded to the digit immediately after the appearance of a "1", without waiting to see if it was a "9". His performance was suggestive of an inability to delay, particularly once he had been primed to respond. Following a complete diagnostic evaluation the youngster was classified as having ADD with Hyperactivity and was placed on a moderate dose of stimulant medication. His academic programming was geared more toward accuracy than speed, and he received resource help to encourage him to modulate his response style. Follow-up contact indicated substantial improvement.

CASE #2

Kevin, a 10-year old boy, was referred for a psychological evaluation by both his school and pediatrician. According to all concerned, Kevin was consistently impulsive, disruptive, inattentive, hyperactive and underachieving. Judging from rating scales and case-history information, his behavior met all the DSM-III criteria and the more stringent set of markers proposed by Barkley. The age of onset for his problems was before 5 years 11 months, his symptoms were chronic and pervasive, his IQ was well over 70, and ratings of his behavior by teachers on standardized checklists were beyond the 5th percentile. His school had demonstrated ample patience with his misbehavior to the extent that they allowed him to leave the classroom and run the hallways for five minutes when he could no longer contain his energy. At the time of the referral, the school was reaching the end of their capacity to deal with Kevin.

Kevin was administered the Delay Task, Standard Vigilance Task, and the Vigilance Task with Distractors. The Delay Task Efficiency Ratio of .75 was in the Borderline Range for his age. The most noteworthy feature of his performance was the very significant degree of variability across the four Time Blocks. At times he inhibited perfectly while at other times his performance approached the Abnormal Range. This marked inconsistency in the context of a profile which shows evidence of the capacity for delay is pattern found often in children whose problems have a strong emotional component. Unlike the more typical ADD profile in which there is little if any evidence of successful inhibition, the emotionally disturbed child's protocol will tend to show clear swings across the session between adequate and inadequate coping.

Kevin's performance on the standard Vigilance Task also fell in the Borderline Range (albeit the near normal end) for both Correct and Commissions. Again, there were indications of inconsistency of performance in that the middle Time Block, unlike the other two, was normal. As is so often the case, Kevin's behavior during the session was telling. Upon missing the "1/9" combination, he would slap himself in the face and complain that the numbers flashed too quickly for him. The examiner noted that Kevin became so anxious about performing on the task that he had trouble concentrating. His discomfort increased on the Distractibility version of the Vigilance Task in which his performance in all respects was in the Borderline Range.

The GDS evaluation indicated that Kevin, at times, could be more impulsive and inattentive than age-mates. However, his poor control did not consistently reach an abnormal range. In fact, he demonstrated a capacity to delay and attend adequately but had difficulty maintaining good performance for reasons that seemed, at least in part, related to emotional issues. His behavior during the testing was of a boy who became unusually anxious and self-denigrating when he met up with frustration or failure.

On the basis of this kind of GDS protocol, a complete psychodiagnostic evaluation is typically suggested. In Kevin's case, a full battery of psychologic tests was administered (WISC-R, WRAT, PIAT, Figure Drawings, TAT, Rorschach, Bender. etc.), in addition to clinical interviews with the boy, his family and teachers. The overall conclusion of this extensive evaluation was that Kevin's problems of self-control were secondary to a severe and chronic emotional disturbance. The onset of his hyperactivity at age 5 coincided with the gunshot murder of his mother. Because his biological father was unavailable (Kevin was born out-of-wedlock), Kevin was raised by his maternal grandparents. The grandparents were very devoted to the boy but, because of their own fears, suspicions and ambivalence about their custody of Kevin, they seriously limited his contacts with others. They also engaged in what appeared to be bizarre rituals around the memory of the boy's mother. This family history as well as a host of other factors left Kevin with intense fears and a dependent, hostile relationship with his grandparents. As he grew older he found it increasingly difficult to manage his anxieties and anger. He would become so flooded by stress that he often would have difficulty organizing himself for even simple tasks. Intensive psychotherapy both with Kevin and his grandparents was initiated.

CASE 3

A seven year old girl was referred to a child development clinic by her teacher for poor school performance, not following instructions, missing assignments, and difficulty with reading. The teacher viewed much of this youngster's behavior as willful and oppositional, and was at her wit's end as to how to help the girl learn. Her parents reiterated the teachers complaints, and added that they felt the child to be distant and aloof much of the time. During the clinical interview this youngster was pleasant and cooperative, but gave the impression of either wanting to be somewhere else or of simply daydreaming.

On the Delay Task she obtained an overall Efficiency Ratio of .82, well within the Normal range. She earned 46 Total rewards. Her responses appeared controlled, orderly, and goal directed. Obviously her difficulties did not lie in the area of impulse control. On the Vigilance Task her difficulties became quite clear. Although she produced just one error of Commission, which is well within normal limits, she made 33 errors of Omission, scoring well beyond the mean on this measure.

Although her motivation throughout the task was quite good, she was unable to maintain the degree of preparedness required to perform effectively on this part of the task. Therefore, we were able to identify her main difficulty as a deficit in sustained attention in the absence of impulsivity, which fits the DSM III classification of Attention Deficit Disorder without Hyperactivity.

Following a more extensive evaluation, she was placed on a very small dose of stimulant medication, and her teacher was informed of the results of the testing. Intervention strategies were offered for both home and school, and her adjustment improved in a satisfactory manner.

Case 4

James, a 13 year-old boy of average intelligence, was referred for GDS testing from an adolescent inpatient unit where he was undergoing treatment for a range of psychiatric problems. The precipitant of this second admission to the unit was the boy's attempt to choke himself with a pencil. James had a history of firesetting, suicidal ideation, poor peer relationships, harming animals, and academic underachievement. Full psychological evaluations had pointed to severe pathology across domains of functioning. He was described as anxious, impulsive, aggressive, depresssed and distractable. In addition to conduct problems, James was considered hyperactive and, upon occassion, had received brief trials of stimulant medication.

Prior to admission, James had resided in a foster home, his fifth placement in as many years. He became unusually unmanageable upon the arrival of a younger foster brother. James had last seen his biological mother three years prior to admission.

While James was considered to be highly impulsive and inattentive, the extent to which those symptoms were secondary to pervasive social and emotional deficits was unclear. Also, there was a need to evaluate treatment with stimulant medication because previous attempts had not been systematically monitored.

James was evaluated with the GDS while off, on, then off medication. (Unfortunately, a double-blind, placebo approach could not be implemented.) James' performance while free of medication fell solidly in the Abnormal Range for both tasks. He demonstrated a clear inability to sustain attention and delay. The practioner noted that, while the scores were generally typical of the ADD/Hyperactive youngster, the test behaviors were somewhat uncharacteristic. Unlike many ADD/Hyperactive children, James appeared cooperative, anxious, eager to please, and concerned about his performance.

After a week of treatment with methylphenidate, James was retested and his scores showed marked improvement. His performances on the Delay and Vigilance Tasks were at or near the Normal range, as he demonstrated improved self-control and attentiveness. While he still displayed considerable immaturity, James nonetheless appeared calmer and more focused.

Two weeks later, James was tested once more, this time free of stimulant medication. Even though he boasted that he was an "expert" at these tests, his performance fell back to baseline levels. He was again unable to exert self-control on the Delay Task, and had a high number of commissive errors on the Vigilance Task.

It was suggested that attention deficits were embedded in a constellation of other difficulties, and that attentional aspects of his problems were responsive to pharmacotherapy. The recommendation was made to continue stimulant medication therapy as part of a comprehensive treatment program.

CONCLUSION

The use of computerized assessment in the evaluation of children referred for ADD/Hyperactivity can make for a more accurate, efficient diagnostic process. It allows the practitioner to incorporate data based on the child's actual behavior. Clinicians who pursue this approach need to select a procedure that is practical, rugged, well-normed, and energetically supported by research. A legitimate concern is that these computer-generated tasks present a relatively narrow range of demands, stimuli, and response modalities. While the ADD child's impulsivity is viewed as generally pervasive across areas of functioning, no security exists that performance on select computerized tasks will necessarily reflect the overall level of impulsivity for a particular child. In all situations, these objective data must be integrated into a comprehensive evaluation of the sort syggested by Barkley (1981, 1985). As part of a sophisticated clinical protocol, a computerized attention task can significantly enhance the diagnostic process.

References

Atkinson, A.W., Cohen, P.C. & Kelly, P.C. (1985, April). Attention deficit disorder: The effects of ritalin on self-esteem a comparison of ACTeRS teacher scale, Conner's parent scale, and Gordon Diagnostic system in diagnosis and managment. Paper presented at the American Academy of Pediatrics Meeting, Atlant, Georgia.

Barkley, R.A. (1981). Hyperactive children: A handbook for diagnosis and treatment. New York: The Guildford Press.

Barkley, R.A. (1985). Development of a multi-method clinical protocol for assessing stimulant drug responding in ADD children. Manuscipt submitted for publication.

Conners, C.K. (1980). Continuous performance test [Computer program]. Kensington, Maryland: Behavioral Medicine, Inc.

Gordon, M. (1979). The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive children. Journal of Abnormal Psychology, 7, 317-326.

Gordon, M. (1982). The Gordon Diagnostic System. .Littleton, Colorado: Clinical Diagnostics, Inc.

Gordon, M. & McClure, F.D. (1983). [The assessment of ADD/Hyperactivity in a public school population]. Unpublished raw data.

Gordon, M. (1984,1985). ADD/Hyperactivity Newsletter. (Available from Clinical Diagnostics, Inc., Educational Services, 300 East Mineral Avenue, #6, Littleton, Colorado 80122).

Greenberg, L.M. (October,1985). An objective measure of response to methylphenidate: Clinical validation of the VIRTEST. Paper presented at the 32nd Annual Meeting of the American Academy of Child Psychiatry, San Antonio, Texas.

Klee, S.H. & Garfinkel, B.D. (1983). The computerized continuous performance task: A new measure of inattention. Journal of Abnormal Child Psychology, 11(4), 489-496.

McClure,F.D. & Gordon, M. (1984). The performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. Journal of Abnormal Child Psychology, 12(4), 561-572.

Rapport, M.D., Dupaul, G.J., Kelly, K. & Jones, J. (1985). Comparing classroom and clinic measures of ADD-H: Dose response effects . Manuscript submitted for publication.

Rosvold, H.E., Mirsky, A.F., Sarason, I., Bransome, E.D. & Beck, L.H. (1956). A continuous performance test of brain damage. Journal of Consulting Psychology, 20, 343-350.