1986

Issues in the Diagnosis of Attention Deficit Disorder: A Cautionary Note on the Gordon Diagnostic System

Richard Milich
University of Kentucky, richard.milich@uky.edu

William E. Pelham
Western Psychiatric Institute and Clinic

Stephen P. Hinshaw
University of California - Berkeley

Issues in the Diagnosis of Attention Deficit Disorder: A Cautionary Note on the Gordon Diagnostic System

The Forum

With this issue we inaugurate a new feature, **The Forum**, which will appear twice a year in issues 2 and 4. Its goal is to provide an opportunity for discussion on issues that are topical and new in psychopharmacology. Consider submitting comments on papers that have appeared in the *Bulletin* (e.g., the paper by Dr. Richard Milich and colleagues presented here). Raise questions that may stimulate responses from others. Material for **The Forum** should be submitted to:

The Forum
Psychopharmacology Bulletin
5600 Fishers Lane, Room 10C-06
Rockville, MD 20857 USA

Submissions will be reviewed by the *Bulletin* Associate Editors. We hope **The Forum** will be a place for lively debate.

—Nina R. Schooler, Ph.D.
Editor-in-Chief

Issues in the Diagnosis of Attention Deficit Disorder: A Cautionary Note on the Gordon Diagnostic System

Richard Milich, Ph.D.,[1] William E. Pelham, Ph.D.,[2] and Stephen P. Hinshaw, Ph.D.[3]

Introduction

In a recent paper Gordon (1986) has compared the strengths and weaknesses of traditional microcomputer-based assessment of attention deficit disorder (ADD) with those offered by the Gordon Diagnostic System (GDS; Gordon & McClure 1983; 1984). Not surprisingly, although careful not to overstate the case, he finds the GDS to be an improvement over the use of standard microcomputers. Nevertheless, several of his criticisms of microcomputers seem forced, and he tends to downplay some of the more serious concerns associated with the GDS. Given that the GDS is receiving widespread attention, with nationwide marketing procedures underway, it seems fitting to examine its strengths and weaknesses independently as compared to more traditional methods of assessment of attention deficit disorder. To facilitate this goal the history of the GDS will be reviewed, since in its relatively brief life several important events have already occurred.

In 1979 the first paper by Gordon in the area of ADD appeared. He presented the results of a study comparing 20 hyperactive

[1]Department of Psychology, University of Kentucky, Lexington, KY.
[2]Western Psychiatric Institute and Clinic, University of Pittsburgh School of Medicine, Pittsburgh, PA.
[3]Department of Psychology, University of California, Berkeley, CA.
Reprint requests: Richard Milich, Ph.D., Department of Psychology, University of Kentucky, Lexington, KY 40506.

and 20 clinic-referred nonhyperactive children on a differential reinforcement of low-rate behavior (DRL) task. The results of this study were quite promising and, as Milich and Kramer (1984) note, offered an intriguing way in which to define and measure the problems in impulsivity exhibited by ADD children. Approximately 3 to 4 years following the publication of this first article, but before the appearance of any subsequent articles, a commercially marketed version of the GDS was made available (Clinical Diagnostics, Inc., 300 E. Mineral Ave., Littleton, CO 80122). Admittedly, Gordon was in the process of collecting large scale normative data, and within the last year an article has appeared replicating the results of the first study (Gordon & McClure 1984). Nevertheless, it does not seem too rash to state that the appearance of the GDS as a commercially available product was premature.

In its original format the GDS contained only the DRL. Halfway through the standardization process a continuous performance test (CPT), modeled after the one originally produced by Rosvold and associates (1956), was included. In its current format the GDS sells for $1295 (originally $2500). For an additional $150 a data analysis program compatible with either the Apple II or IBM-PC microcomputer is available. For an additional $18 a newsletter is available, published four times a year, that deals with attention deficit disorder in general and the GDS specifically.

Argument

Although Gordon (1986) discusses the flexibility of the GDS, basically the DRL and the CPT are the only two tasks that can be run on the machine. The question raised is whether the information obtained justifies the relatively high cost of the apparatus. It is important to keep in mind that for the same expenditure of funds one could purchase a relatively sophisticated microcomputer system with printer and two disc drives, at least. Such a system would give one markedly improved flexibility, in terms of software available for assessment procedures, word processing capabilities, and data analysis programs, to name just several of the common uses for microcomputers.

As noted, purchasing the Gordon Diagnostic System involves a relatively major financial investment, money that could perhaps be spent in a more cost-efficient fashion. Therefore, the question becomes whether the GDS is supplying valuable information that cannot be obtained in a more economical manner. To answer this question it may be helpful to examine the research available on the GDS. We will summarize briefly the research for the two different tasks: the Delay task, as the DRL is now called, and the Vigilance task, as the CPT task is named.

As noted, two published studies exist to date supporting the validity of the Delay task in differentiating, with a relatively high rate of accuracy, ADD and non-ADD clinic-referred children. However, as Gordon (1985) himself notes, the validity of an assessment instrument can be no greater than the measures employed to establish the criterion groups against which the new instrument is being validated. Thus, for example, when the criterion groups (ADD vs. non-ADD) were established by meeting cutoffs on three rating scales and one laboratory test (the Matching Familiar Figures Test), the Delay task was found to agree with the criterion measures on 15 of 16 cases (94%). When the children were classified as ADD or not by meeting criteria on three of the measures, agreement with the GDS was found in 29 of 32 cases (91%). Although such results certainly offer support for the sensitivity of the GDS in identifying ADD children, they also indicate that four relatively inexpensive and easily obtained measures can do a comparable job of identification.

There are two other aspects of the Delay task that are somewhat disquieting when considering its role in the assessment of ADD. First, the normative data (*see* Gordon & McClure 1983) have not found gender differences in performance on the Delay task. Such a finding contrasts markedly with the accepted prevalence figures for ADD, which range anywhere from 4:1 to 10:1, boys to girls. Whereas Gordon (1985) suggests that there may be many girls with significant deficits in attention who are being overlooked, the Delay task is one of the few measures in the ADD field that fails to find sex differences.

Furthermore, the Delay task has not consistently been shown to be drug sensitive. For

example, Barkley (1985) found none of the three measures from the Delay task (rewards, responses, efficiency ratio) to differentiate among placebo and two dosages of methylphenidate (.3 and .5 mg/kg). In contrast, most of the other measures he collected, including parent and teacher ratings, playroom observations, and the Vigilance task, all showed evidence of drug sensitivity. Considering that most measures employed in the assessment of ADD routinely show sensitivity to drug effects, this lack of sensitivity of the Delay task gives one pause as to the utility of this procedure.

Along these lines, it is interesting to note that Gordon (1985), in one of his most recent presentations summarizing the current research on the GDS, focuses entirely on the Vigilance task. His motivation for this decision is not clear but it may well reflect this task's greater sensitivity to stimulant medication. In fact, in the Barkley (1985) study, the Vigilance task was one of the few measures also to differentiate low and high dosages of methylphenidate. It is certainly not surprising that the Vigilance task should exhibit such sensitivity; these procedures have been around for 30 years, and a wealth of studies have documented their drug sensitivity and ability to differentiate ADD and non-ADD children (*see* Douglas 1983). The questions then become whether the GDS Vigilance task offers a significant improvement over more conventional CPT procedures and whether its validity and utility justify its great expense. To answer these questions we will examine Gordon's (1986) criticisms of conventional microcomputer assessments.

Gordon (1986) notes basically three limitations of conventional microcomputers in the assessment of ADD: their size and cumbersomeness, the unreliability of administration, and their delicacy. It is obvious that the first and third points deal with issues of practicality and utility, rather than validity per se. Although at one time the GDS may have been more mobile and hence practical than portable microcomputers, the ones currently available are no more cumbersome, and considerably less expensive, than the GDS. For example, an Apple II-C, which is small and easily portable, can be obtained for approximately $700, and a number of relatively inexpensive CPT tasks compatible with the II-C are available in the public domain. Similarly, Gordon's third point regarding the delicacy of the machines does not seem to be a real issue. Altogether we have administered a variety of information processing tasks on Apple II microcomputers to several hundred ADD children, with no unusual problems and no damage to hardware or software.

Thus, the only substantive issue regarding microcomputer assessment raised by Gordon (1986) concerns the reliability of the evaluation. Here Gordon means that different systems present the stimuli in slightly different fashions and may require different responses (e.g., pressing a spacebar vs. a single letter). In one sense, Gordon is correct in that different CPT programs or microcomputers may produce slightly different response patterns, making it difficult to establish specific cutoff or criterion scores to arrive at a formal diagnosis. The GDS, which presents the same stimuli in a standardized fashion to anyone who uses the machine, could generate such criteria.

First, however, the issue of norms is crucial to such an endeavor. The standardization sample reported by Gordon and McClure (1984) contains 700 boys and girls from ages 3 through 16. Thus, at any given age and gender level, there may be fewer than 30 subjects. In addition, the cutoff scores presented by Gordon and McClure are not separated by gender, a problem noted earlier regarding the GDS. Thus, the viability of these norms for diagnostic use may be questionable.

Second, as Gordon (1986) himself notes, a diagnosis should not be based upon a single score on a single instrument. Instead, it must be derived from the convergence of multiple sources of data. From that perspective, the score on the GDS is, by itself, no more diagnostically significant than that obtained from teacher rating scales or systematic classroom observations. In fact, one could argue that the presentation of formalized normative data and cutoff scores may inadvertently encourage clinicians to rely primarily or even solely on such a measure, at the expense of other information. This is a realistic possibility, given the lack of standardized criteria in the field and given particularly the designation of "Diagnostic System" to the hardware and procedures of the GDS. The problems inherent in single-instrument diagnoses have been noted

by Ullmann and associates (1985) with respect to the use of cutoff scores from the Conners (1973) Abbreviated Teacher Rating Scale to diagnose ADD. There is no reason to suspect that a single score from a continuous performance test will escape such problems.

Furthermore, although performance on a CPT task is currently thought to be useful in defining and thus diagnosing ADD, there is sufficient disagreement among professionals regarding the precise nature of the attention deficit in ADD children (*see* Douglas 1983) that currently popular versions of the CPT task, including Gordon's, may not be thought relevant in future diagnostic systems. If one uses a microcomputer to administer an attention task, one is able to acquire new software when accepted diagnostic tasks change. The GDS, however, lacks such flexibility.

Conclusion

The GDS, especially the Vigilance task, appears to have some utility in aiding with the diagnosis of ADD. However, legitimate concerns can be raised regarding the cost efficiency of purchasing such an apparatus, especially when apparently comparable information can be obtained at considerably less expense. Further, the apparent simplification of the diagnostic process may predispose clinicians to resort to simplistic assessment procedures. Finally, we feel the need to reiterate our concern regarding the aggressive nature in which the GDS has been marketed, especially given the relatively meager data available when the apparatus was first offered for sale and the inherent limitations of any single instrument for the purpose of diagnosis.

References

Barkley, R.A. "Assessment of stimulant drug responding in ADD-H children." Paper presented at the 93rd Annual Meeting of the American Psychological Association, Los Angeles, CA, August 1985.

Conners, C.K. Rating scales for use in drug studies with children. *Psychopharmacol. Bull.* (Special Issue, *Pharmacotherapy of Children*), 24–84, 1973.

Douglas, V.I. Attentional and cognitive problems. In: Rutter, M., ed. *Developmental Neuropsychiatry*. New York: Guilford Press, 1983. pp. 280–329.

Gordon, M. The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive children. *J. Abnorm. Child Psychol.* 7:317–326, 1979.

Gordon, M. Microprocessor-based assessment of attention deficit disorders. *Psychopharmacol. Bull.* 22:288–290, 1986.

Gordon, M. "Current GDS research: The vicissitudes of validation." Paper presented at the 93rd Annual Meeting of the American Psychological Association, Los Angeles, CA, August 1985.

Gordon, M., and McClure, F.D. "The objective assessment of attention deficit disorder." Paper presented at the 91st Annual Meeting of the American Psychological Association, Anaheim, CA, August 1983.

Gordon, M., and McClure, F.D. "Assessment of attention deficit disorders using the Gordon Diagnostic System." Paper presented at the 92nd Annual Meeting of the American Psychological Association, Toronto, Canada, August 1984.

Milich, R., and Kramer, J. Reflections on impulsivity: An empirical investigation of impulsivity as a construct. In: Gadow, K.D., ed. *Advances in Learning and Behavioral Disabilities*. Greenwich, CT: JAI Press, 1984. pp. 57–94.

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

Ullmann, R.K.; Sleator, E.K.; and Sprague, R.L. A change of mind: The Conners Abbreviated Teacher Rating Scales reconsidered. *J. Abnorm. Child Psychol.* 13:567–577, 1985.