Journal of Abnormal Child Psychology, Vol. 29, No. 3, 2001, pp. 215–228

The Ecological Validity of Delay Aversion and Response Inhibition as Measures of Impulsivity in AD/HD: A Supplement to the NIMH Multimodal Treatment Study of AD/HD

Mary V. Solanto,[1,10] Howard Abikoff,[2] Edmund Sonuga-Barke,[3] Russell Schachar,[4] Gordon D. Logan,[5] Tim Wigal,[6] Lily Hechtman,[7] Stephen Hinshaw,[8] and Elihu Turkel,[9]

Received September 1, 1999; revision received February 8, 2000; accepted April 21, 2000

Impulsivity is a primary symptom of the combined type of Attention Deficit/Hyperactivity Disorder (AD/HD). The Stop Signal Paradigm is premised upon a primary deficit in inhibitory control in AD/HD, whereas the Delay Aversion Hypothesis, by contrast, conceptualizes impulsivity in AD/HD, not as an *inability* to inhibit a response, but rather as a *choice* to avoid delay. This study compared the ecological validity of the Stop Signal Task (SST) and Choice-Delay Task (C-DT) measure of delay aversion, with respect to their relative utility in discriminating AD/HD children from normal control participants, and their correlations with classroom observations and with ratings of impulsivity and other core AD/HD symptoms on the Conners and SNAP-IV checklists. The tasks exhibited modest discriminant validity when used individually and excellent discriminant validity when used in combination. The C-DT correlated with teacher ratings of impulsivity, hyperactivity, and conduct problems, and with observations of gross motor activity, physical aggression, and an AD/HD composite score. The SST correlated with the observations only. These results suggest that delay aversion is associated with a broad range of AD/HD characteristics whereas inhibitory failure seems to tap a more discrete dimension of executive control.

KEY WORDS: AD/HD children; impulsivity; inhibitory control; Stop Signal Task; delay aversion.

Impulsivity is one of the core features of the most commonly recognized subtype of Attention Deficit/Hyperactivity Disorder (AD/HD). An impulsive response may be defined as one that is executed with insufficient forethought, planning, or control, and is therefore inaccurate or maladaptive. Beyond this broad definition, however, there is little consensus with respect to specific criteria for, or conceptual modeling of, this behavioral construct.

Behaviors that may be considered "impulsive" comprise a topographically diverse set. Examples include (a) responding before instructions are given or before a

[1]Division of Child and Adolescent Psychiatry, Schneider Children's Hospital, Long Island Jewish Medical Center, New Hyde Park, New York; currently at Division of Child Psychiatry, Mt. Sinai Medical Center, New York.
[2]Division of Child and Adolescent Psychiatry, Schneider Children's Hospital, Long Island Jewish Medical Center, New Hyde Park, New York; currently at New York University Child Study Center, New York, New York.
[3]Department of Psychology, University of Southampton, Southampton, UK.
[4]Department of Child Psychiatry, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada.
[5]Department of Psychology, University of Illinois, Champaign, Illinois.
[6]Department of Pediatrics, University of California, Irvine, California.
[7]Division of Child Psychiatry, McGill University, Montreal, Quebec, Canada.
[8]Department of Psychology, University of California, Berkeley, California.
[9]Department of Psychiatry, Hillside Hospital, Long Island Jewish Medical Center, New Hyde Park, New York.
[10]Address all correspondence to Dr. Mary Solanto-Gardner, Division of Child Psychiatry, Mt. Sinai Medical Center, Box 1230, One Gustav Levy Place, New York, New York 10029-6574; e-mail: mary_solanto-gardner@mssm.edu.

0091-0627/01/0600-0215$19.50/0 © 2001 Plenum Publishing Corporation

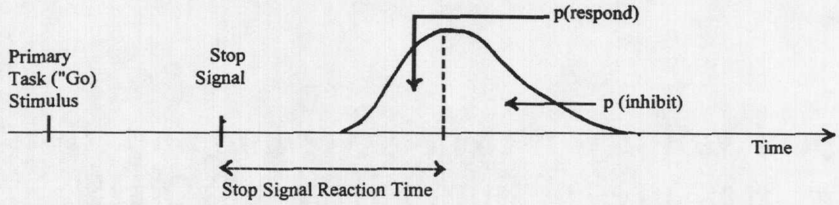

Fig. 1. Time sequence of stimuli on Stop Signal Task, also showing hypothetical distribution of reaction times.

question is completed, (b) responding without first considering all response options, (c) failing to withhold a motor or cognitive response to an irrelevant or inappropriate stimulus, and (d) acting before considering the consequences of a socially offensive or aggressive behavior. Numerous laboratory tasks have been developed in efforts to operationalize and measure these and other impulsive behaviors. Among those that have been used to assess impulsivity in AD/HD children are the Matching Familiar Figures Test (MFFT; Kagan, 1965), Porteus Mazes (Palkes, Stewart, & Kahana, 1968), Continuous Performance Test (Losier, McGrath, & Klein, 1996), Stroop Color-Word Interference Test (Barkley, Grodzinsky, & DuPaul, 1992), Draw-A-Line Slowly and Draw-A-Line Fast tests (Levy & Hobbes, 1979), and the competitive game developed by Atkins (Atkins & Stoff, 1993; Atkins, Stoff, Osborne, & Brown, 1993) to differentiate between hostile/impulsive and instrumental aggressive acts.

Perhaps predictably, given the diversity of processes that may be involved, scores on these and other purported laboratory measures (LMs) of impulsivity correlate poorly both among themselves, and with parent and teacher ratings of impulsive behaviors in natural settings (Barkley, 1991; Milich & Kramer, 1985). More recent research (Kindlon, Mezzacappa, & Earls 1995) suggests that impulsivity may be differentially expressed in motor, cognitive, social, and emotional domains. This heterogeneity may account for the low correlations among tests and measurements in prior impulsivity research.

Multiple, nonmutually exclusive explanatory models have been proffered for the underlying cognitive or neuropsychological deficit(s) that give rise to impulsivity. In recent years, much of this theoretical debate has focused on the extent to which deficits in inhibitory control are implicated in the impulsive behaviors displayed by children with AD/HD. On the one hand, some argue (e.g. Barkley & Biederman, 1997) that generalized deficits of this sort are at the heart of this disorder. The most compelling evidence for this view comes from research using the Stop Signal Paradigm, developed by Logan et al. (Logan, Cowan, & Davis, 1984; Logan & Cowan, 1984) and first applied to childhood behavior disorders by Schachar and Logan (1990). This paradigm provides an index of the AD/HD child's ability to inhibit a prepared motor response. Such an ability might be expressed in such everyday behaviors as checking a swing at a bad pitch in baseball, or stopping oneself from running out into the street to continue chasing an errant ball. On the Stop Signal Task (SST; (Fig. 1) the individual performs a visual choice reaction time task (the primary task). On a proportion of trials, randomly selected, a tone ("stop signal") is presented *after* the primary task stimulus ("go-signal") and is the cue to the individual to *inhibit* the response to the go-signal on that trial. According to the model developed by Logan and colleagues (Logan et al., 1984), the probability of inhibiting the response depends on the outcome of a race between the "go" process and the stopping process. If the go process is faster than the stopping process, the individual emits the response; if the stopping process is faster, the response is inhibited. The speed of the stopping process, the stop signal reaction time (SSRT), may be inferred on the basis of the distribution of reaction times on trials without a stop-signal, and the probability of inhibition.

The interval between the stop signal and the individual's own mean reaction time (mRT) is systematically varied across trials on the SST. Extensive prior research with this paradigm (Logan & Cowan, 1984) has shown that the probability of successful inhibition is a direct function of the length of this stopping interval: longer intervals are associated with a greater probability of inhibition. Superior inhibitory control is manifested in a higher and steeper curve for the probability of inhibition plotted against the length of the stopping interval. Children with pervasive ADDH (*DSM-III*) or pervasive AD/HD (*DSM-III-R*) were shown to have both a flatter slope of inhibition (Schachar & Logan, 1990; Schachar, Tannock, Marriott, & Logan, 1995) and a longer mean SSRT (Schachar et al., 1995) than did normal children. In addition, methylphenidate was found to enhance the probability of inhibition and reduce SSRT in children with AD/HD (Tannock, Schachar, Carr, Chajczyk, & Logan, 1989; Tannock, Schachar, & Logan, 1995).

On the other hand, other theories take a motivational approach to the disorder. These accounts do not regard AD/HD as the result of disinhibitory psychopathology but rather as the expression of an altered motivational state that leads to an altered response to reinforcement parameters (especially magnitude and delay). Among these are that children with AD/HD have a reduced sensitivity to reinforcement such that more immediate, frequent, or intense rewards are required to maintain appropriate performance and behavior (Barkley, 1989; Haenlein & Caul, 1987); are *overly* responsive to immediate rewards (Douglas & Parry, 1994; Rapport, Tucker, DuPaul, Merlo, & Stoner, 1986); are less able to delay gratification or resist temptation (Mischel, Shoda, & Rodriguez, 1989); or are higher in "stimulation-seeking" behaviors, which are needed to compensate for inherently low levels of central nervous system arousal in AD/HD (Zentall & Meyer, 1987; Zentall & Zentall, 1983).

The Delay Aversion Hypothesis, developed by Sonuga-Barke and colleagues (Sonuga-Barke, Taylor, Sembi, & Smith, 1992b), characterizes impulsive behavior not as the consequence of a relative *inability* to inhibit a response, but rather as the result of a rational *choice* to avoid delay, which the individual finds aversive. This view was first tested in a paradigm ("Choice-Delay Task" for purposes of the present paper) in which the child repeatedly chooses between a large reward, which is associated with a period of delay, and a small reward, which is not associated with a delay. The basic choice offered in this task is similar to that used in previous research on delay of gratification (Mischel et al., 1989). However, a key feature of this task is that after the choice for the delayed option is made the participant cannot switch to the small immediate reward during the trial. This means that the preference for the large reward does not involve the inhibition of the response for the small reward, but rather the active initial choice of an alternative.

In their first paper (Sonuga-Barke et al., 1992b), Sonuga-Barke and colleagues compared the performance of community identified pervasively hyperactive children and normal control children in two experiments in which children chose between a small reward (1 point) associated with a delay of 2 s and a large reward (2 points) associated with a delay of 30 s. This choice was made across four different conditions. Hyperactive children did not differ from normal control children on three of these, which are as follows: (a) When rewards were not followed by delay so that the immediate reward led to shorter sessions and more reward over sessions, both groups preferred the immediate reward. (b) When rewards were followed by delay so that both alternatives were associated with the same session length but the delayed reward was associated with more reward overall, both groups showed preference for the large delayed reward. (c) When there was no postreward delay and both alternatives had the same session length ("time constraint" condition), so that the immediate reward was associated with more reward over the session, both groups chose the small immediate reward. However, when there were a limited number of trials on which to choose ("trials constraint") so that the small reward was associated with shorter sessions but less reward overall, the hyperactive children showed a significantly smaller mean preference for the large delayed reward (18%) than did the control children (48%). This pattern of results suggested that overall delay (as reflected in session length), rather than prereward delay or reward size, was the key motivating factor that produced impulsive responding. This experimental evidence further indicates that choice between delayed and immediate rewards in the context of a trials constraint provides the most useful measure of delay sensitivity, given its ability to discriminate between AD/HD and control children.

Interestingly, there was no evidence that reward parameters (magnitude or rate) per se exerted particular control over the hyperactive children's behavior. This is consistent with prior research, which has not yielded any consistent evidence that children with AD/HD differ systematically from normal children in valuation or sensitivity to magnitude of reinforcement (Douglas & Parry, 1994; Pelham, Milich, & Walker, 1986; Solanto, 1990, 2001). However, it is important to recognize that, while emphasizing the salience of delay reduction to hyperactive children, this paradigm does not discount the role of reward size altogether. All children will be sensitive to changes in reward size to some extent and it is unlikely that hyperactive children's aversion to delay overrode this sensitivity completely. It would be compatible with the delay aversion hypothesis to postulate that one could reduce the preference for the small immediate reward (and less delay) under the trials constraint condition by increasing the size of the large delayed reward to an extent that countered the dominant motivation to avoid delay. It is an empirical question as to whether a greater increase in reward magnitude would be required to bring about the same shift in preference for hyperactive children than for their nonhyperactive counterparts.

In subsequent studies, Sonuga-Barke and colleagues investigated the extent to which delay aversion may account for the findings of poorer performance by children with AD/HD on traditional measures of attention, memory, and impulsivity. Aversion to delay was related to longer serial reaction times (Sonuga-Barke & Taylor, 1992), shorter self-selected stimulus exposure times, and worse memory recognition (Sonuga-Barke, Taylor,

& Heptinstall, 1992a), as well as shorter latency to response on the MFFT (Sonuga-Barke, Houlberg, & Hall, 1994) in hyperactive as compared to normal children. However, delay aversion did *not* account for the higher frequency of errors on the MFFT in the hyperactive group (Sonuga-Barke, Williams, Hall, & Saxton, 1996), suggesting that factors such as meta-cognitive strategies (Douglas, 1979) or executive functions (Barkley, 1997) also play a role.

The foregoing indicates that both the Stop Signal and Delay Aversion Paradigms have demonstrated utility in the characterization of deficits in AD/HD. However, little is known about the specific role played by the psychological processes tapped by these tasks (response inhibition/delay aversion) in the manifestation of specific symptoms. In this paper we will ask the following question: To what extent are delay aversion and deficits in response inhibition implicated in the impulsive behaviors displayed by children with AD/HD? To put it another way, how valid are the SST and the Choice-Delay Task (C-DT) as models of the sorts of impulsive behavior displayed by children with AD/HD? In trying to answer this question it is of particular value to explore the associations between task performance (e.g. SSRT on the SST and choice for the delayed reward on the C-DT) and direct classroom observations and parent and teacher ratings. Notwithstanding that there is no established, gold standard index of impulsivity, standardized questionnaires and objective observations yield "real-world" behavioral measures of impulsivity that are not only face-valid, but have demonstrated discriminant validity and sensitivity to treatment effects as well. With the exception of a single study of the SST by Pliszka (Pliszka, Borcherding, Spratley, Leon, & Irick, 1997), the ecological validity of these tasks with respect to impulsivity in the real-world settings of home and school has not been investigated.

Barkley (1991) has described ecological validity as the extent to which performance on a LM represents the actual behaviors of interest as they occur in natural settings. It is thus a type of predictive validity. Barkley further enumerates ways that ecological validity may be assessed: (1) differences between patient and control groups on the LM; (2) correlations between the LM and the ecological criterion, which may be caregiver ratings or direct observations; (3) correlations between the LM and another LM of the same construct already shown to have ecological validity; and (4) the demonstration that the LM shows similar changes when exposed to experimental manipulations (e.g. stimulant medication) known to affect the ecological criterion. The current study employs both of the first two types of ecological validation. Furthermore, with respect to Method 2, we employ caregiver ratings as well as direct observations, which Barkley recommended, given the respective liabilities of each of these measures.

The main purpose of the following study, therefore, was to compare the ecological validity of the SST and the C-DT as measures of impulsive behavior in children with AD/HD. We utilized the extensive behavioral database of rigorously diagnosed children with AD/HD participating in the NIMH-sponsored Multimodal Treatment Study of AD/HD (MTA;[11] Arnold et al., 1997). We assumed that the tasks measure an underlying construct of impulsivity that may be manifested behaviorally in diverse ways. There were two stages to the study: In the first, we examined the validity of measures derived from the two tasks in discriminating between AD/HD cases and noncases and calculated values for their sensitivity and specificity. In the second stage, we examined correlations between each task score and measures from the MTA database, that we hypothesized *a priori* would be most highly related to the construct of impulsivity: *DSM-IV* Impulsivity ratings on the SNAP-IV; Conners Hyperkinesis Index (HI) score; and scores on the Interference (Composite) category of the Classroom Observation Code (COC) that has face validity as a measure of "impulsive" behavior. In addition, we examined correlations between the task scores and the other core symptoms of AD/HD and the associated oppositional, conduct, and aggressive behavior problems, as assessed on the SNAP-IV, Conners Rating Scale, and the COC. We predicted that the C-DT and the SST scores would each correlate more highly with scales or factors assessing impulsivity than with other core or associated symptoms of AD/HD. We further postulated that the

[11] MTA Cooperative Group for Treatment Phase (through 14 months): The MTA is a cooperative treatment study performed by six independent research teams in collaboration with the staff of the Division of Clinical and Treatment Research of the National Institute of Mental Health (NIMH), Rockville, Maryland and the Office of Special Education Programs (OSEP) of the U.S. Department of Education (DOE). The NIMH Principal Collaborators are Peter S. Jensen, MD, L. Eugene Arnold, MEd, MD, John E. Richters, PhD, Joanne B. Severe, MS, Donald Vereen, MD, and Benedetto Vitiello, MD. Principal Investigators and Coinvestigators from the six sites are as follows: University of California at Berkeley/San Francisco (UO1 MH50461): Stephen P. Hinshaw, PhD, Glen R. Elliott, MD, PhD; Duke University Medical Center (UO1 MH50447): C. Keith Conners, PhD, Karen C. Wells, PhD, John S. March, MD, MPH; University of California at Irvine/Los Angelos (UO1 MH50440): James M. Swanson, PhD; Dennis P. Cantwell, MD; Timothy Wigal, PhD; Long Island Jewish Medical Center/Montreal Children's Hospital (UO1 MH50453): Howard B. Abikoff, PhD, Lily Hechtman, MD; New York State Psychiatric Institute/Columbia University/Mount Sinai Medical Center (UO1 MH50454): Laurence L. Greenhill, MD, Jeffery H. Newcorn, MD; University of Pittsburgh (UO1 MH50467): William E. Pelham, PhD, Betsy Hoza, PhD. Helena C. Kraemer, PhD (Stanford University), is statistical and design consultant. The OSEP/DOE Principal Collaborator is Ellen Schiller, PhD.

relative strength of these correlations for the two tasks would give an indication of the relative validity of the corresponding theoretical models of impulsive behavior in AD/HD. Finally, given controversy about whether the combination of AD/HD and oppositional defiant disorder (ODD) or conduct disorder (CD) represents a distinct subtype of AD/HD (Schachar & Tannock, 1995), as well as marked inconsistencies in the results of prior studies of the impact of comorbid ODD/CD on SST scores (Schachar & Logan, 1990; Schachar & Tannock, 1995) and the complete lack of data on the effect of comorbid disorders on the C-DT task, we compared subgroups of AD/HD children with and without comorbid ODD/CD.

METHOD

Participants

Children newly enrolled in the MTA study at the Long Island Jewish Medical Center (LIJMC), Montreal and Berkeley sites were invited to participate. As participants in the MTA study, all children met selection criteria for the parent study, as follows: age between 7.0 and 9.9 years at enrollment (children may have been as much as a year older by the time of their participation in the current supplemental study); grade between first and fourth; diagnosis of AD/HD (combined type) on the basis of the Diagnostic Interview for Children and Adolescents (DISC 3.0) conducted with the parent (Shaffer et al., 1996); a score at least 1.0 $SD >$ norm on the Parent Conners Hyperkinesis Index *or* Parent Conners Hyperactivity Factor; a score at least 1.5 $SD >$ norm on the Teacher Conners Hyperkinesis Index *or* Conners Teacher Hyperactivity Factor *or* IOWA Teacher Conners Inattention/Overactivity factor; WISC-III Full Scale, Verbal or Performance IQ ≥ 80 *or* Scales of Independent Behavior ≥ 80; absence of Bipolar Disorder, Psychosis, Pervasive Developmental Disorder, severe Obsessive–Compulsive Disorder, Tourette or chronic serious tics; absence of major neurological or medical illness. Separate informed consent to participate in the supplemental study was obtained from the parent and informed assent from the child. All children were tested in the current study at the conclusion of their evaluation for the parent study and before any treatment had been initiated; no children were receiving medication at the time of their participation in the current study. A total of 77 children, 86% of whom were male, received the SST. Of these, 45 children, 90% of whom were male, received the C-DT as well, the others having been tested before the current C-DT revision was ready for use. Demographic and behavioral characteristics of participants are displayed in Table I. The socioeconomic status represented in the AD/HD sample was broad; gross annual family income ranged from less than $10,000 to over $75,000, averaging $40,000–$50,000. There were no significant differences across sites within this sample on these measures.

The children in the normal control group were recruited from two sources: Two elementary schools, in the same school districts from which participants in the LIJMC MTA sample were drawn, circulated a letter home to parents describing the study and asking for volunteers. Ratings were obtained from teachers on the Conners Teacher Questionnaire concerning those children who volunteered to confirm the absence of significant behavioral, learning, or emotional problems. Children were required to have a score within one standard deviation of the normative mean for age and sex on the Conners Hyperkinesis Index on this questionnaire.

Children were also drawn from the longitudinal normal control group participating in the MTA study at the Montreal site. Only those children whose ages were within the range represented in the AD/HD sample participating in the current supplemental study were selected for inclusion. The final normal control group of 29 children did not differ from the AD/HD sample with respect to

Table I. Demographic and Diagnostic Characteristics of Sample

	CHOICE- DELAY ($n = 45$)[a]		STOP SIGNAL ($n = 77$)[a]		CONTROL ($n = 29$)	
	Mean (*SD*)	Range	Mean (*SD*)	Range	Mean (*SD*)	Range
Age (years)	8.45 (0.87)	7.27–10.3	8.48 (0.92)	7.07–10.6	8.7 (0.88)	7.62–9.98
IQ(WISC-III)[b]	97.7 (14.3)	76–131	98.55 (13.71)	71–131	Not available	
Conners Index[c]						
Parent	19.70 (5.79)	6–30	19.71 (5.79)	6–30	Not available	
Teacher	19.03 (4.97)	11–28	18.45 (5.36)	0–28	5.7 (6.32)	0–22

[a]All participants who received the Choice-Delay Task also received the Stop Signal Task.
[b]Correlations between IQ and each task score were negligible.
[c]Conners 10-item Hyperkinesis Index.

age (Table I) or gender distribution (83% male). Informed consent to participate was obtained from the parent and informed assent from the child. The normal control children performed the SST and the C-DT only; COC data was not available for these children.

Procedures

Tasks

Testing for the supplemental study for the AD/HD group was conducted at the conclusion of the MTA baseline assessment visit, after other measures had been completed. Children in the normal control group were tested at their schools. Tasks were administered on a computer. The SST was presented first, followed by the C-DT. The choice to administer the SST first was made because this task is more effortful and less enjoyable for most children than is the C-DT, and thus was more likely to be adversely affected by fatigue or boredom if administered later in the session. Pilot work with the C-DT indicated that it was relatively robust vis-a-vis order effects. Given these task characteristics, a fixed rather than counterbalanced order, with SST first, was selected in order to minimize overall variance in task performance.

Stop Signal Task. The SST used in this study was programmed in MEL (Micro Experimental Lab). The task consisted of six blocks, and with scripted instructions (available from the first author), required approximately 30 min to complete. The first block (32 trials) was the visual reaction time task, in which an "O" or an "X" appeared in the middle of the screen for 1000 msec. X's and O's were upper-case letters, 4.0 mm high and 2.0 mm wide. The child was instructed to press the correspondingly labeled key as quickly as possible when each letter appeared. (Stickers labeled "X" or "O" were affixed to the "4" and "6" keys, respectively.) Only those responses occurring within 1000 msec of stimulus onset were recorded. X's and O's were presented with equal frequency within each block in random order. Offset of the primary task stimulus was followed by a wait-interval of 1500 msec and then by presentation of a fixation point in the form of a small cross (2 × 2 mm) for 500 msec, followed by onset of the next letter stimulus.

During Blocks 2 through 6 (48 trials each), the same task was presented with the following modification: an auditory stop signal (500 Hz, 150 msec) was presented through a standard laptop or desktop speaker on 16 (33%) randomly selected trials and was the cue to the child to *refrain* from responding to the primary task stimulus on that trial. The stop-signal was presented with equal frequency at four different "stopping intervals" before the individual's own expected primary task reaction time: 100, 250, 350, and 500 msec. The reaction time used in this calculation was the child's mRT for correct hits during the immediately preceding block. Thus, the actual interval between the onset of the primary task stimulus and the stop signal could assume any of four different values for each subject within each trial block: mRT minus 100, mRT minus 250, mRT minus 350, and mRT minus 500. The stop-signal was most often presented after the onset of the primary task stimulus. In those instances in which mRT for the previous trial block was 0 (no hits) or mRT was *less than* the stopping interval, the stop signal and primary task stimulus were presented *simultaneously*. The stop signal was *never* presented *before* the onset of the primary task stimulus. In order to discourage children from slowing down to await a possible beep, each child was reinstructed at the start of each trial block to respond as quickly as possible to the letter stimulus while also inhibiting a response if the stop signal was heard. The second block of trials was considered a "practice" stopping block; data only from Blocks 3–6 were entered into the analysis.

Calculation of Stop Signal Response Indices. SSRT was calculated using the method previously employed by Schachar and colleagues (Schachar et al., 1995), in which a distribution of reaction times on go-signal trials is created (Fig. 1). This distribution is integrated until the integral is equal to the probability of response in the presence of the stop signal [p(respond | signal)]. The corresponding point on the x-axis represents the point in time at which the stopping process was finished. It is assumed that responses with latencies shorter than this reaction time are not inhibited, whereas responses with longer latencies are appropriately inhibited. Alternately stated, when the reaction time on a given trial is less than this critical reaction time, the "go-process" will win the "race," whereas when the reaction time on a given trial is longer than this critical value the stopping process will be faster and will win the "race," resulting in appropriate inhibition of the response. The point at which the two processes are equivalent — that is the reaction time that cuts off a proportion of the distribution that is equal to [p(respond | signal)] — is assumed to represent the latency of the stopping process from the onset of the "go-signal."

In the current study, SSRT was estimated as follows:

1. The probability of response in the presence of the stop signal [p(respond | signal)] was calculated for each of the four stopping intervals (100, 250, 350, and 500 msec). Each of these calculations included a correction for nonresponse due to failures to respond (omission errors) on "go-signal" trials, as follows"

$$[p(\text{respond} \mid \text{signal})] = \frac{x - \text{correct rejections}}{x - xy}$$

where x, the number of stop-signal trials at each stopping interval, is 16; correct rejections are the number of correct inhibitions of response on stop-signal trials; and y is the probability of omission on go-signal trials in Blocks 3–6.

2. The reaction times for hits on go-signal trials for Blocks 3–6 (including both accurate and inaccurate responses to X's and O's) are rank-ordered. The reaction time at the "nth" percentile is identified, where "n" corresponds to p(respond | signal). This reaction time is the latency from the onset of the go-signal, which approximates the latency between the go-signal and the conclusion of the stopping process. Subtracting out the stop-signal delay (i.e. the interval between the onset of the go-signal and the onset of the stop-signal; Fig. 1) yields the reaction time to the stop signal for each stopping interval. SSRT is the average of SSRTs across stopping intervals.

Other indices derived from the SST were as follows: HitRT was the mean reaction time for hits; p(inhibition) was the probability of correctly inhibiting a response when the stop signal was presented; p(omission) was the probability of failing to respond on go-signal trials.

Choice-Delay Task. A C-DT written expressly for this study was presented using scripted instructions (available from the author) and required approximately 30 min to complete. The child used the mouse to choose between two rectangles, each measuring 6.5 cm × 3.6 cm, presented side by side on the computer screen: a green square labeled "1 point" and a blue square labeled "2 points." The instructions explained to the children that they were about to play a game in which they could earn points, and that each point earned would be exchanged for a nickel at the end of the game. Children practiced manipulating the mouse to highlight and select one of the two boxes. In five subsequent practice trials they were coached to choose alternating boxes and then were asked to compare the difference in the waiting periods. The relative difference in waiting period was confirmed by the examiner. If, at the end of this practice there was any doubt as to whether the child thoroughly understood the operation of the task, particularly the difference in delay depending on choice of reward, another set of 3–5 practice trials was administered. Before the test trials, the child was instructed that he/she would have 20 "tries" on which to earn points, that one chip would be placed on a grid for each try so that he/she would always know how may tries were left, and that there was no time limit so he/she could take as much time as he/she wanted. Children and families received no other monetary compensation for participating in the study.

Task parameters were the same as those used in the trials constraint condition of Experiment 2 in the original paper (Sonuga-Barke et al., 1992b). Choices of the 1-point and 2-point rewards were followed by a "prereward" delay of 2 s or 30 s, respectively, before the number of points earned in that trial were posted on the screen. One-point versus 2-point reward choices and postreward delays were chosen because these were the parameters used in the experiment that demonstrated that delay aversion was more important than reward immediacy in determining choice, and also because pilot work before the current study indicated that the 1:2 point ratio was most effective in avoiding floor and ceiling effects and maximizing differences between groups. Also replicating the method of the original paper, two blocks of 20 trials were presented; at the start of each trial a bingo chip was placed on a 20-block grid to indicate the number of trials remaining; the child was given his/her monetary reward after each of the two trial blocks. The side of presentation of the large reward was counterbalanced within participants' handedness.

Independent Measures

SNAP-IV Scale. This is a checklist, completed by parents and teachers, of *DSM-IV* symptom criteria for AD/HD and ODD. Responses are weighted from 0 to 3 (*Not at all* to *Very Much*) and summed (Swanson & Carlson, 1994).

Classroom Observation Code (COC). The COC has been shown to discriminate the classroom behavior of hyperactive and normal children (Abikoff, Gittelman, & Klein, 1980; Abikoff, Gittelman-Klein, & Klein, 1977). Using this measure (Abikoff & Gittelman, 1985), the classroom behavior of all children participating in the MTA was observed and coded within 12 mutually exclusive off-task, motor, and acting-out categories. For purposes of the current study, several categories were selected as best operationalizing the core symptoms of AD/HD: the Interference and Interference to Teacher categories were selected and summed to generate a composite measure of impulsive behavior; the Off-Task category was selected to operationalize inattention; and the Gross Motor-Standing and Gross Motor-Vigorous (GMV) categories were summed to yield Gross Motor-All as a measure of hyperactivity. The Interference category includes the following operationalized subcategories: interruption of other students, production of sounds, annoying behavior, and clowning. Interference to Teacher includes interfering behaviors directed specifically at the teacher, such as leaving seat and going up to the teacher or calling out to the teacher. The Off-Task category monitors behaviors in which the child, after initiating the appropriate task-relevant behavior, attends to stimuli other than the assigned work. GMV includes vigorous, unexpected gross motor movements, such as suddenly swinging between desks on the way back to one's seat. Gross Motor-Standing refers to motor activity that results in the child leaving his seat and standing. An AD/HD composite score was calculated by summing the Off-Task, composite Interference score, and Gross Motor-All scores. The AD/HD composite has been used as an outcome measure in treatment studies of AD/HD (Gittelman-Klein et al., 1980). The category of physical aggression was defined to include forceful movement directed by the child at another person (by using either his body or using a material object as an extension of the hand), as well as destruction of another person's property.

Observations were conducted during teacher-led lessons or independent academic seatwork under teacher supervision or both. During a scheduled observation period, two children were observed in each classroom—the AD/HD study child, and a comparison child (not a participant in the current study) of the same gender and ethnicity identified by the teacher as of average comportment. The comparison child was observed as a control for the level of activity and expectations for deportment in the particular classroom. The observers were blind to the diagnostic status of the children. Observation sessions lasted 32 min, during which time the children were observed in alternating 4-min blocks, divided into 16 continuous 15-s intervals, yielding 16 min of data on each child. For all categories selected for this study, with the exception of Off-Task, a modified time sampling strategy was used such that only the first occurrence of the behavior in the 15-s interval was scored. For Off-Task, a timed criterion was required and the behavior was scored if it occurred throughout the entire 15-s interval. The observation scores represent the mean percentage of intervals in which the behaviors occurred. The score utilized in this study was the observation score for the AD/HD child minus that for the corresponding comparison child.

Determination of interobserver agreement for the behavioral categories was conducted on approximately 10% of the classroom observations in the entire MTA sample, using phi coefficients as an index of reliability of interval scores. The phi coefficient is a correlational measure of reliability that ranges from $+1$ to -1 and, like kappa, does not overestimate the degree of agreement between observers. For data in a 2×2 table, phi "provides nearly identical values to kappa" (Gelfand & Hartmann, 1975). The phi coefficients for all categories were in the acceptable range, ranging from 0.80 to 1.00. The mean phi across all observation categories was 0.91.

Conners Parent and Teacher Questionnaires. The Conners Questionnaires have been widely used in studies of AD/HD. The original 93-item Parent and 39-item Teacher forms were used in the MTA study (Conners, 1990).

Co-morbid Diagnoses. The DISC-IV, administered to the parents of children in the AD/HD sample, was used to diagnose comorbid ODD, CD, and Anxiety Disorder. For purposes of this study, comorbid Anxiety Disorder referred to diagnoses *other than* Simple Phobias, and included Generalized Anxiety Disorder, Social Phobia, Obsessive–Compulsive Disorder, and Separation Anxiety. Because a higher percentage of children with ODD/CD had an additional comorbid Anxiety Disorder than did those without ODD/CD (37% vs. 20%, respectively) each child's self-rating on the Multidimensional Anxiety Scale for Children (MASC; March, Parker, Sullivan, Stallings, & Conners, 1997), as well as age, was used as covariates in these comparisons.

RESULTS

Stop Signal Task

Results on the SST were analyzed after first excluding data from 14 children with a high probability of omissions on "go-trials" (greater than 20%), and 10 children with a very low probability of inhibition (less than 10%) on stop-signal trials. Three children were members of both these subgroups. Thus, 21 children were omitted, leaving a final sample of 56. Children with a high percentage of omissions were apparently not attending to the task for a significant proportion of time. Their nonresponses on stop-signal trials therefore most likely do not represent successful inhibitions, but rather failure to perform the task according to instructions. In cases in which inhibition of response occurs on less than 10% of stop-signal trials, SSRTs are computed on the basis of fewer than two values per stopping interval, yielding artifactually long values for SSRT. Performance on SST parameters are detailed in Table II, and are comparable to results reported for other AD/HD samples (Schachar et al., 1995), illustrating the expected increase in probability of inhibition with increases in the stopping interval. Comparisons with the control group reveal robust mean differences in SSRT and probability of inhibition (Table II).

Table II. *T*-tests of Mean Differences Between AD/HD and Control Groups for SST and C-DT Variables

Score	Group	Mean	*SD*	t^a	*p*(one-tailed)
SST: *p*(omit)	AD/HD	0.06	0.05	1.01	.159
	Control	0.05	0.05		
SST: *p*(inhibit)	AD/HD	0.41	0.14	−3.99	.000
	Control	0.54	0.15		
SST: SSRT	AD/HD	435.73	280.89	3.52	.001
	Control	290.16	104.04		
SST: HitRT	AD/HD	763.97	186.45	−0.17	.434
	Control	769.44	119.23		
C-DT: Large reward (% choice for)	AD/HD	0.34	0.26	−3.71	.000
	Control	0.58	0.27		

[a]Equal variances not assumed.

Delay Aversion Paradigm

On the C-DT, children with AD/HD chose the large reward on 34% ($SD = 26\%$) of trials in Block 1. Choice of the large reward increased to 49% ($SD = 34\%$) on Block 2, presumably influenced by the dispensing of reinforcement between blocks, which is likely to have increased the incentive value of the points administered during the task. The control group increased their choice of the delayed reward from 58% ($SD = 27\%$) to 69% ($SD = 34\%$). A 2×2 repeated-measures analysis of variance (ANOVA) with two groups and two time points revealed a significant effect of group ($F = 12.32$, $p = .001$) and of block ($F = 6.65$, $p = .012$) with no interaction between group and block ($F = .36$, *ns*). Because of the significant improvement from Block 1 to Block 2, only the score from Block 1 will be used in subsequent analyses.

Correlations Among Task Scores

Neither age nor IQ was significantly correlated with task scores. Therefore neither was used as a covariate in the analyses. Correlations among the C-DT score and the four SST scores were examined, with the following results: SSRT correlated highly and negatively with probability of inhibition ($r = -.805$, $p < .05$), and also correlated with probability of omission ($r = .239$, $p < .05$), but did not correlate with Hit ("Go") Reaction Time (HitRT). Probability of inhibition correlated with HitRT ($r = .247$, $p < .05$), but not with probability of omission. Probability of omission correlated only with HitRT ($r = .286$, $p < .05$). C-DT correlated only with HitRT ($p = -.349$, $p < .05$). (All *p*-values are one-tailed.) Thus, a faster stopping time was correlated, as expected, with a greater likelihood of inhibition. A faster reaction time to the go-signal, conversely, was associated with a reduced likelihood of appropriately inhibiting a response and also with a reduced likelihood of omitting a response to the go-signal. A higher probability of choice for the delayed large reward was associated with a faster reaction time to the go-signal.

Discriminant Validity in Comparisons With Control Group

Separate discriminant function analyses were performed using each of the SST and C-DT variables in order to determine the percentage of individuals correctly assigned to the AD/HD group (sensitivity) or the control group (specificity) on the basis of each of these measures. In addition, the possibility that the discriminant validity might be enhanced by using more than one index simultaneously was explored by entering the best predictors together in stepwise fashion in the same discriminant function. Despite the robust group mean differences for SST and C-DT scores by *t*-tests, these measures discriminated modestly between groups (Table III). Furthermore, entering *p*(inhibit) and SSRT together in the same equation did not enhance the discriminant validity of *p*(inhibit) alone. Probability of choice for the large (delayed) reward displayed better sensitivity but worse specificity than the comparable values for SST indices, yielding an overall correct percent that was approximately equal. Entering C-DT and *p*(inhibit) together yielded marked increases in specificity, sensitivity, and overall correct classification rates of between 85% and 90%.

Correlations Between Task Performance and Behavioral Measures

Descriptive statistics for the selected behavioral measures are presented in Table IV, and Pearson *r* correlations

Table III. Specificity and Sensitivity of SST and C-DT Variables

Index	Sensitivity (%)	Specificity (%)	Overall correct (%)
SST: *p*(omit)	44.1	72.4	53.41
SST: *p*(inhibit)	66.1	72.4	68.18
SST: SSRT	50.8	82.8	61.36
SST: Hit RT	57.6	44.8	53.41
C-DT:Large reward (% choice for)	76.9	64.3	71.64
C-DT and *p*(inhibit) (stepwise)	89.3	85.0	87.50

(one-tailed, uncorrected alpha) between SST and C-DT scores and the behavioral factors are displayed in Table V. The results indicated that on the SST, there were significant correlations between the SSRT and three measures from the COC: Interference-composite, AD/HD composite, and Physical Aggression. For probability of inhibition, the pattern of significant correlations (not shown) was identical, which was expected, given the high correlation between the two indices. For the C-DT, the frequency of choice of the large reward correlated significantly in the expected negative direction with SNAP-IV Teacher Impulsivity and Hyperactivity ratings as well as with the Conners Hyperkinesis Index and Conduct Problem scales. In addition, the C-DT correlated significantly and negatively with the COC categories of Interference-composite, Gross Motor-All, AD/HD composite, and Physical Aggression. Among a total of 11 correlations for each of the two tasks, there were 10 significant correlations—3 for the SST and 7 for the C-DT; *one* of these correlations would be significant by chance alone.

Table IV. Descriptive Statistics for the SST and C-DT Test Samples

	SSRT ($n = 56$)[a]	C-DT ($n = 45$)[b]
SNAP-IV (*DSM-IV*)-Teacher		
Inattention	18.81 (6.42)	21.08 (5.27)
Impulsivity	5.81 (2.82)	5.09 (2.98)
Hyperactivity	11.36 (4.35)	11.66 (4.11)
Oppositional Defiant Disorder	10.59 (6.82)	9.70 (6.78)
Conners Questionnaire-Teacher		
Hyperkinesis Index	18.62 (5.45)	19.03 (4.97)
Conduct problem	14.68 (9.37)	14.54 (9.67)
Classroom Observation Code[c]		
Off-task	0.07 (0.12)	0.06 (0.14)
Interference-composite	0.14 (0.14)	0.13 (0.14)
Gross Motor-all	0.0097 (0.029)	0.0015 (0.047)
AD/HD composite	0.22 (0.19)	0.19 (0.28)
Physical aggression	0.0036 (0.011)	0.0058 (.014)

Note. Values represent Mean (*SD*).
[a]Data were missing on at least one scale for six of these cases.
[b]Data were missing on at least one scale for three of these cases.
[c]These scores represent the difference between target and comparison children in percent of intervals in which the behavior occurred. See text for details.

Impact of Comorbidity

In the SST sample, 71% of participants had a comorbid ODD/CD or Anxiety Disorder (other than Simple Phobia): 7% of the sample had an Anxiety Disorder; 40% had ODD/CD; and 24% had both disorders. The corresponding figures for the C-DT sample were 65%, 11%, 30%, and 24%, respectively. Comparison on SST performance (SSRT, probability of omission, probability of inhibition, and HitRT) between subgroups of the AD/HD

Table V. Pearson Correlations Between Task Scores and Behavioral Factors

	SSRT ($n = 56$)[a]	C-DT ($n = 45$)[b]
SNAP-IV (*DSM-IV*)-Teacher		
Inattention	*ns*	*ns*
Impulsivity	*ns*	−.378**
Hyperactivity	*ns*	−.278*
Oppositional Defiant Disorder	*ns*	NS
Conners Questionnaire-Teacher		
Hyperkinesis Index	*ns*	−.309*
Conduct problem	*ns*	−.265*
Classroom Observation Code[c]		
Off-task	*ns*	*ns*
Interference-composite	.298*	−.375**
Gross Motor-all	*ns*	−.355**
AD/HD composite	.265*	−.280*
Physical aggression	.306*	−.395**

Note. There were no significant correlations between either task and ratings by parents on the SNAP-IV scale or the Conners Questionnaire; *ns* = nonsignificant.
[a]Data were missing on at least one scale for six of these cases.
[b]Data were missing on at least one scale for three of these cases.
[c]These scores represent the difference between target and comparison children in percent of intervals in which the behavior occurred. See text for details.
$^{*}p < .05$, one-tailed. $^{**}p < .01$, one-tailed.

sample with ODD/CD ($n = 38$) and without ODD/CD ($n = 20$), using MASC score as a covariate, yielded no significant differences for any SST index. Similarly, there was no difference in C-DT performance between children with ODD/CD ($n = 20$) and those without ($n = 18$).

DISCUSSION

In comparisons between children with AD/HD and normal control participants, there were robust group mean differences for SSRT and probability of inhibition on the SST, as well as for the C-DT index. In classifying approximately 70% of cases correctly these two measures individually [compared favorably to] standard neuropsychological batteries tapping executive functions (Grodzinsky & Barkley, 1999; Schreiber, Javorsky, Robinson, & Stern, 1999) and performed better than did some commercially marketed diagnostic aides (Rielly, Cunningham, Richards, Elbard, & Mahoney, 1999). These results suggest that both delay aversion and inhibitory failure are implicated in AD/HD. Furthermore, when the C-DT score and p(inhibit) were considered together as discriminant predictors, sensitivity, specificity, and overall correct classification exceeded 85%. If these findings are replicated in new, larger samples of children with ADHD and normal control children, these measures may prove a useful addition to diagnostic batteries. It is important, however, to note that SST data from 27% of the participants were omitted due to questionable validity of these scores, which may limit the utility and generalizability of SST scores.

Performance on the C-DT was significantly improved after dispensing the monetary reward. This observation not only suggests that *both* reward salience and delay aversion can modify performance on the C-DT, but also points to the potential *malleability* of lessened tolerance for delay in children with AD/HD by manipulation of behavioral contingencies.

In general, correlations between task scores and the behavioral variables were modest (all $r < .4$). However, it is important to bear in mind that selection of children with AD/HD naturally markedly truncated the range of scores represented in the data set and thus reduced the magnitude of the correlations. Furthermore, these results compare favorably with those reviewed by Barkley (1991), who concluded that among measures used to assess sustained attention, omission errors on the continuous performance test (CPT) and Children's Checking Task exhibited the greatest ecological validity on the basis of "moderate" correlations with teacher ratings of attention (.21–.51 and .44–.61 for these two tasks, respectively). Among measures of impulsivity, errors of commission on the CPT also showed modest correlations, ranging from .32 to .44, with teacher ratings of impulsivity and hyperactivity. Notably, in these studies, scores from AD/HD and normal children were *combined* in the calculation of correlations, expanding the range of scores and thereby increasing the possible magnitude of the correlations.

Current results revealed that preference for the large reward on the C-DT was correlated (inversely) with impulsivity as rated by teachers, as well as with teacher ratings of hyperactivity and conduct problems, and direct observations of gross motor activity, physical aggression, and the AD/HD composite measure. SSRT and probability of inhibition on the Stopping Task were associated only with direct observations of impulsivity (Interference-composite), physical aggression, and the AD/HD composite score. Replicating results for other LMs of AD/HD symptoms (Barkley, 1991), neither task displayed any correlation with parent ratings.

The findings of task correlations with measures of hyperactivity as well as impulsivity corroborate previous research showing that impulsivity and hyperactivity share a common variance in children with AD/HD (Lahey et al., 1988), and also support the view that the relative inability or disinclination to delay a response is a core deficit underlying and unifying the symptomatic manifestations of the combined type of AD/HD (Barkley, 1997). The more extensive range of correlations between C-DT performance and these diverse behavioral measures suggest that the C-DT may be of more general significance than is the SST. In this context, it is interesting to note that Sonuga-Barke (1994) has argued that inattention and overactivity as well as impulsivity, can be seen as expressions of delay aversion. Although formally distinct clinical manifestations, they can be viewed as functionally equivalent, their function being the escape, avoidance, or at least reduction of delay. These functional properties of impulsiveness may be relatively easily identified in those situations where choices between more and less delayed alternatives are available. But what of situations in which children have no choice between different delays and where escape or avoidance of delay is impossible? How would delay aversion be manifest in these situations? Starting from the large body of experimental evidence on the effects of attention and action on time perception (Anderson, Burd, Dodd, & Kelk, 1980; Falk & Bindra, 1954; Fraisse, 1963; Hicks, Miller, Gaes, & Bierman, 1977; Kikkawa, 1983; Langer, Wadner, & Werner, 1961), Sonuga-Barke (1994) has argued that hyperactive children's inattention and overactivity represent attempts to reduce the perceived (but not the actual) length of time in passing. For example, it has been shown that when an individual's attention is directed toward aspects of a situation that reduce awareness of the passage of time (nontemporal stimuli) and away from aspects that emphasize it (temporal stimuli), the length of that time period seems shorter (Brown, 1985; Hicks, Miller, & Kinsbourne, 1976). There is also evidence that activity can serve the same function. It follows from this that effectively, if not objectively, delay can be reduced, under fixed delay conditions, by maximizing attention to nontemporal stimulation, fidgeting, or otherwise eliciting stimulation from the environment. Patterns of attention and action that would serve to reduce the subjective experience of time in passing correspond closely with the types of dysregulated activity and sustained attention problems seen as characterizing the attentional style associated with AD/HD.

Compared with the C-DT, the SST exhibited more limited patterns of relationship with the dependent variables. SSRT was correlated only with direct observations on the COC and not with any of the teacher or parent ratings of attention, impulsivity, or hyperactivity. These findings, as well as the notable lack of correlation between C-DT and SST scores, suggest that these two paradigms tap different components of the AD/HD phenotype. *Stopping* may be a discrete component of executive control that is not necessarily highly correlated with other forms of self-regulation and that may be dependent upon unique neural mechanisms. Further research is necessary to validate the distinctions between the self-control processes operationalized in the SST and C-DT paradigms and to investigate their relative merits for characterizing the cognitive and behavioral deficits in AD/HD.

Comorbidity with a second disruptive behavior disorder (ODD or CD) did not significantly affect scores on either task. The absence of an effect of comorbid ODD/CD is consistent with the results of a recent meta-analysis of eight studies showing no differences on the SST between children with AD/HD and those with AD/HD plus CD (Oosterlaan, Logan, & Sergeant, 1998).

ACKNOWLEDGMENTS

The authors extend their sincere appreciation to the following: Martin Hall at the University of Southampton for development of the C-DT software; Lori Feirman for yeoman work in testing participants at the LIJMC site and compiling data from all sites; Martha Carlin for pilot-testing at LIJMC; Josee Champagne and Carrie Dandy for testing at Montreal, and Joe Garcia for testing at Berkeley; and Karen Fleiss, Joy Etcovitch, and Joe Garcia for coordination at the LIJMC, Montreal, and Berkeley sites, respectively. We are also indebted to Dr. Fred Jaffe, psychologist, and Nancy Lindenauer, principal, for the participation of the Searingtown School on Long Island, and also to Michael Pagliuca, District Administrator of Special Education, for the participation of New York Public Schools District 26. We thank the children themselves for participating, and also their parents and teachers for completing the rating scales. The authors extend appreciation to Rosemary Tannock, PhD, and Jeffrey Newcorn, MD, for their insightful comments on a previous version of this paper.

REFERENCES

Abikoff, H., & Gittelman, R. (1985). Classroom Observation Code: A modification of the Stony Brook Code. *Psychopharmacology Bulletin, 21,* 901–909.

Abikoff, H., Gittelman, R., & Klein, D. F. (1980). A classroom observation code for hyperactive children: A replication of validity. *Journal of Consulting and Clinical Psychology, 48,* 555–565.

Abikoff, H., Gittelman-Klein, R., & Klein, D. F. (1977). Validation of a classroom observation code for hyperactive children. *Journal of Consulting and Clinical Psychology, 45,* 772–783.

Anderson, B., Burd, L., Dodd, J., & Kelk, K. (1980). A comparative study of time estimation. *Journal of American Indian Education, 19,* 1–4.

Arnold, L. E., Abikoff, H. B., Cantwell, D. P., Conners, C. K., Elliott, G. R., Greenhill, L. L., Hechtman, L., Hinshaw, S. P., Hoza, B., Jensen, P. S., Kraemer, H. C., March, J. S., Vereen, D., & Well, K. C. (1997). NIMH collaborative multimodal treatment study of children with ADHD (MTA): II. Design, methodology, and protocol evolution. *Archives of General Psychiatry, 54,* 865–870.

Atkins, M. S., & Stoff, D. M. (1993). Instrumental and hostile aggression in childhood disruptive behavior disorders. *Journal of Abnormal Child Psychology, 21,* 165–178.

Atkins, M. S., Stoff, D., Osborne, M. L., & Brown, K. (1993). Distinguishing instrumental and hostile aggression: Does it make a difference? *Journal of Abnormal Child Psychology, 21,* 355–365.

Barkley, R. A. (1989). The problem of stimulus control and rule-governed behavior in attention deficit disorder with hyperactivity. In L. M. Bloomingdale & J. Swanson (Eds.), *Attention Deficit Disorder: Current concepts and emerging trends in attentional and behavioral disorders of childhood* (pp. 203–232). Elmsford, NY: Pergamon Press.

Barkley, R. A. (1991). The ecological validity of laboratory and analogue assessments of ADHD symptoms. *Journal of Abnormal Child Psychology, 19,* 149–178.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121,* 65–94.

Barkley, R. A., & Biederman, J. (1997). Toward a broader definition of the age-of-onset criterion for attention-deficit hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 1204–1210.

Barkley, R. A., Grodzinsky, G., & DuPaul, G. J. (1992). Frontal lobe functions in attention deficit disorder with and without hyperactivity: A review and research report. *Journal of Abnormal Child Psychology, 20,* 163–188.

Brown, S. W. (1985). Time perception and attention: The effects of prospective versus retrospective paradigms and task demands in period estimation. *Perception and Psychophysics, 38,* 115–124.

Conners, C. K. (1990). *Manual for the Conners' Rating Scales.* Toronto, Ontario: Multihealth Systems.

Douglas, V. I. (1979). Toward a clearer definition of the attentional deficit of hyperactive children. In G. A. Hale & M. Lewis (Eds.), *Attention and the Development of Cognitive Skills* (pp. 173–247). New York: Plenum Press.

Douglas, V. I., & Parry, P. A. (1994). Effects of reward and nonreward on frustration and attention in attention deficit disorder. *Journal of Abnormal Child Psychology, 22,* 281–302.

Falk, J. L., & Bindra, D. (1954). Judgment of time as a function of serial position and stress. *Journal of Experimental Psychology, 47,* 279–282.

Fraisse, P. (1963). *The psychology of time.* New York: Harper & Row.

Gelfand, D. M., & Hartmann, D. P. (1975). *Child behavior analysis and therapy.* New York: Pergamon Press.

Gittelman-Klein, R., Abikoff, H., Pollack, E., Klein, D. F., Katz, S., & Mattes, J. (1980). A controlled trial of behavior modification and methylphenidate in hyperactive children. In C. Whalen & B. Henker (Eds.), *Hyperactive children: The social ecolory of identification and treatment* (pp. 221–243). New York: Academic Press.

Grodzinsky, G. M., & Barkley, R. A. (1999). Predictive power of frontal lobe tests in the diagnosis of attention deficit hyperactivity disorder. *Clinical Neuropsychology, 13,* 11–21.

Haenlein, M., & Caul, W. (1987). Attention deficit disorder with hyperactivity: A specific hypothesis of reward dysfunction. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 356–362.

Hicks, R. E., Miller, G. W., Gaes, G., & Bierman, K. (1977). Concurrent processing demands and the experience of time in passing. *American Journal of Psychology, 90,* 431–446.

Hicks, R. E., Miller, G. W., & Kinsbourne, M. (1976). Prospective and retrospective judgments of time as a function of amount of information processed. *American Journal of Psychology, 89,* 719–730.

Kagan, J. (1965). Reflection-impulsivity and reading ability in primary grade children. *Child Development, 36,* 609–628.

Kikkawa, M. (1983). The effects of distribution of attention on prospective and retrospective estimations of long temporal intervals. *Journal of Child Development, 19,* 14–22.

Kindlon, D., Mezzacappa, E., & Earls, F. (1995). Psychometric properties of impulsivity measures: Temporal stability, validity and factor structure. *Journal of Child Psychology and Psychiatry, 36,* 645–661.

Lahey, B. B., Pelham, W. E., Schaughency, E. A., Atkins, M. S., Murphy, H. A., Hynd, G., Russo, M., Hartdagen, S., & Lorys-Vernon, A. (1988). Dimensions and types of attention deficit disorder. *Journal of the American Academy of the Child and Adolescent Psychiatry, 27,* 330–335.

Langer, J., Wadner, S., & Werner, H. (1961). The effect of danger upon the experience of time. *American Journal of Psychology, 74,* 94–97.

Levy, F., & Hobbes, G. (1979). The influences of social class and sex on sustained attention (vigilance) and motor inhibition in children. *Australia New Zealand Journal of Psychiatry, 13,* 231–234.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91,* 295–327.

Logan, G., Cowan, W., & Davis, K. (1984). On the ability to inhibit simple and choice reaction time responses: A model and a method. *Journal of Experimental Psychology and Human Perceptual Performance, 10,* 276–292.

Losier, B. J., McGrath, P. J., & Klein, R. M. (1996). Error patterns on the continuous performance test in non-medicated and medicated samples of children with and without ADHD: A meta-analytic review. *Journal of Child Psychology and Psychiatry, 37,* 971–987.

March, J., Parker, J., Sullivan, K., Stallings, P., & Conners, C. K. (1997). The Multidimensional Anxiety Scale for Children (MASC): Factor structure, reliability, and validity. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 554–565.

Milich, R., & Kramer, J. (1985). Reflections on impulsivity: An empirical investigation of impulsivity as a construct. In K. Gadow & I. Bialer (Eds.), *Advances in learning and behavioral disabilities* (Vol. 3). Greenwich, CT: JAI.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science, 244,* 933–938.

Oosterlaan, J., Logan, G. D., & Sergeant, J. A. (1998). Response inhibition in AD/HD, CD, co-morbid AD/HD + CD, anxious, and control children; a meta-analysis of studies with the stop task. *Journal of Child Psychology and Psychiatry, 39,* 411–425.

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

Pelham, W. E., Milich, R., & Walker, J. L. (1986). Effects of continuous and partial reinforcement and methylphenidate on learning in children with attention deficit disorder. *Journal of Abnormal Psychology, 95,* 319–325.

Pliszka, S. R., Borcherding, S. H., Spratley, K., Leon, S., & Irick, S. (1997). Measuring inhibitory control in children. *Journal of Developmental and Behavioral Pediatrics, 18,* 254–259.

Rapport, M. D., Tucker, S. B., DuPaul, G. J., Merlo, M., & Stoner, G. (1986). Hyperactivity and frustration: The influence of control over and size of rewards in delaying gratification. *Journal of Abnormal Child Psychology, 14,* 191–204.

Rielly, N. E., Cunningham, C. E., Richards, J. E., Elbard, H. J., & Mahoney, W. J. (1999). Detecting Attention Deficit Hyperactivity Disorder in a communications clinic: Diagnostic utility of the Gordon Diagnostic System. *Journal of Clinical and Experimental Neuropsychology, 21,* 685–700.

Schachar, R., & Logan, G. D. (1990). Impulsivity and inhibitory control in normal development and childhood psychopathology. *Developmental Psychology, 26,* 710–720.

Schachar, R., & Tannock, R. (1995). Test of four hypotheses for the comorbidity of attention-deficit hyperactivity disorder and conduct disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 639–648.

Schachar, R., Tannock, R., Marriott, M., & Logan, G. (1995). Deficient inhibitory control in attention deficit hyperactivity disorder. *Journal of Abnormal Child Psychology, 23,* 411–437.

Schreiber, H. E., Javorsky, D. J., Robinson, J. E., & Stern, R. A. (1999). Rey-Osterrieth Complex Figure performance in adults with attention deficit hyperactivity disorder: A validation study of the Boston Qualitative Scoring System. *Clinical Neuropsychology, 13,* 509–520.

Shaffer, D., Fisher, P., Dulcan, M., Davies, M., Piacentini, J., Schwab-Stone, M., Lahey, B. B., Bourdon, K., Jensen, P. S., Bird, H. R., Canino, G., & Regier, D. (1996). The NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC-2.3): Description, acceptability, prevalence rates, and performance in the MECA study. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 865–877.

Solanto, M. V. (1990). The effects of reinforcement and response-cost on a delayed response task in children with attention deficit hyperactivity disorder: A research note. *Journal of Child Psychology and Psychiatry, 31,* 803–808.

Solanto, M. V. (2001). Attention-Deficit/Hyperactivity Disorder: Clinical Features. In M. V. Solanto, A. F. T. Arnsten, & F. X. Castellanos (Eds.), *Stimulant drugs and ADHD: Basic and clinical neuroscience.* (pp. 3–30). New York: Oxford University Press.

Sonuga-Barke, E. J. S., Houlberg, K. & Hall, M. (1994). When is "impulsiveness" not impulsive? The case of hyperactive children's cognitive style. *Journal of Child Psychology and Psychiatry, 35,* 1247–1253.

Sonuga-Barke, E. J. S., & Taylor, E. (1992). The effect of delay on hyperactive and non-hyperactive children's response times: A research note. *Journal of Child Psychology and Psychiatry, 33,* 1091–1096.

Sonuga-Barke, E. J. S., Taylor, E., & Heptinstall, E. (1992a). Hyperactivity and delay aversion II: The effects of self-imposed and externally imposed presentation time on memory. *Journal of Child Psychology and Psychiatry, 33,* 399–409.

Sonuga-Barke, E. J. S., Taylor, E., Sembi, S., & Smith, J. (1992b). Hyperactivity and delay aversion I: The effect of delay on choice. *Journal of Child Psychology and Psychiatry, 33,* 387–398.

Sonuga-Barke, E. J. S., Williams, E., Hall, M., & Saxton, T. (1996). Hyperactivity and delay aversion III: The effect on cognitive style of imposing delay after errors. *Journal of Child Psychology and Psychiatry, 37,* 189–194.

Swanson, J., & Carlson, C. L. (1994). *DSM-IV rating scale for ADHD and ODD.* Unpublished manuscript.

Tannock, R., Schachar, R. J., Carr, R. P., Chajczyk, D., & Logan, G. D. (1989). Effects of methylphenidate on inhibitory control in hyperactive children. *Journal of Abnormal Child Psychology, 17,* 473–491.

Tannock, R., Schachar, R., & Logan, G. (1995). Methylphenidate and cognitive flexibility: Dissociated dose effects in hyperactive children. *Journal of Abnormal Child Psychology, 23,* 235–266.

Zentall, S. S., & Meyer, M. J. (1987). Self-regulation of stimulation for ADD-H children during reading and vigilance task performance. *Journal of Abnormal Child Psychology, 15,* 519–536.

Zentall, S. S., & Zentall, T. R. (1983). Optimal stimulation: A model of disordered activity and performance in normal and deviant children. *Psychological Bulletin, 94,* 446–471.