Perceptual and Motor Skills, 2012, 114, 1, 59-74.

DEVELOPMENT AND USE OF A COMPUTERIZED TEST, MATH–CPT, TO ASSESS ATTENTION[1]

DUBI LUFI AND NADAV FICHMAN

Department of Behavioral Sciences
Yezreel Valley College

Summary.—The present research describes the development of a new measure of attention, the Mathematics Continuous Performance Test (MATH–CPT), which uses a sequence of simple mathematical questions projected onto a computer screen as visual stimuli. A new approach to testing was developed: it has more complicated stimuli and has an open reaction time allowing participants to react according to individual pace. The development of reliability and validity of the MATH–CPT is described. Discriminant function analysis of 240 normal control participants compared with 63 individuals with ADHD showed correct classification of 91.6% of participants in both groups. The MATH–CPT diagnosed a sample of participants with ADHD better than another CPT-type test, the Test of Variables of Attention. This is an initial step in developing a new measure of attention and to assist with the diagnosis of adolescents and young adults with ADHD.

Attention Deficit Hyperactivity Disorder

Attention Deficit Hyperactivity Disorder (ADHD) is a common childhood syndrome described as "a persistent pattern of inattention and/or hyperactivity-impulsivity that is more frequent and severe than is typically observed in individuals at a comparable level of development" [American Psychological Association (APA), 2000, p. 85]. The syndrome is diagnosed by using various methods such as interviews, rating scales, and reviews of personal histories, medical backgrounds, educational and psychological measures, and neuropsychological tests. There are many facets to ADHD, which makes it very difficult to diagnose. It is now becoming more popular to incorporate various computerized procedures into the assessment of ADHD.

Assessment of ADHD With Continuous Performance Tests

Common computerized procedures to help in the assessment of ADHD are the Continuous Performance Tests (CPT). The idea behind such tests is to have a relatively simple and monotonous stimulus to which a fast response is required. The testee has to process the stimuli and respond quickly by pressing a computer keyboard or a different device attached to the computer. The test lasts a relatively long time, between 10 and 30 minutes, and can be presented visually or auditorily. Each visual stimulus appears for a fraction of a second on the computer screen, and the tes-

[1]Address correspondence to Dubi Lufi, Ph.D., Kibbutz Yifat 36583, Israel or e-mail (dubilufi@gmail.com or dubil@yvc.ac.il).

DOI 10.2466/08.11.22.PMS.114.1.59-74 ISSN 0031-5125

tee is asked to respond as quickly as possible. The main goal of the tests is to assess sustained attention (Pei-Chun, Ching, Yen, Yi, Tzung, I-Hui, *et al.*, 2005). In addition, most of the tests assess attention (the quality of response), impulsivity (uncontrolled fast reactions), speed of processing (reaction time), and variability (consistency in reaction time), depending on the specific test; performed in various clinical groups, such as individuals with ADHD, children with learning disabilities, brain-injured patients, and other clinical populations.

The use of CPT-type tests has proliferated in recent years. Riccio, Reynolds, and Lowe (2001) counted more than 400 studies published on the topic. Halperin, Sharma, Greenblatt, and Schwartz (1991) described the reason for their popularity in the following way: "The recent upsurge in popularity of this test is related to its high face validity, to the availability of microcomputers on which it is frequently programmed and administered, and to the (relatively untested) belief that it objectively assesses attention independent of verbal, perceptual, and other cognitive processing abilities" (p. 603).

Originally, the stimuli used in CPT-type tests used letters of the alphabet in which the testee had to respond only to the appearance of specific letters while ignoring the others. Further developments resulted when researchers developed easier or more complex tests by changing the stimulus from letters to shapes (Cronblatt, Risch, Faris, Friedman, & Erlenmeyer-Kimling, 1988; Greenberg, Kindschi, & Corman, 1999), to numbers (Nuechterlein, Edell, Norris, & Dawson, 1986; Gordon, 1987), or to pictures (Aman, Kern, McGhee, & Arnold, 1993). Others developed tests for preschoolers. Mahone (2005) reviewed nine different CPT-type tests designed for preschoolers. Several of the tests included visual as well as auditory stimuli. In most of the tests, the interpretations of the responses are reported in terms of omissions, commissions, and reaction time (see a discussion by Borgaro, Pogge, DeLuca, Bilginer, Stokes, & Harvey, 2003). None of the tests used an open time format, a format different from the simple recognition of "target-only" or "identical-pairs" paradigms. In open time format, the stimuli changes to the next one only after the testee responds to the problem presented on the screen.

It was reported that the CPT-type tests had the best diagnostic power among various tests used in neuropsychology in diagnosing ADHD (Doyle, Biederman, Seidman, Weber, & Faraone, 2000; Barkley, 2006; Bennett, Zentall, & Franch, 2006; Soreni, Crosbie, Ickowicz, & Schachar, 2009). Research showed that the tests could discriminate well among normal groups and various clinical groups who have ADHD, learning disabilities, conduct disorders, schizophrenia, depression, and the schizotypal personality disorder; discussion about this topic is beyond the scope of this paper.

New Mathematics Continuous Performance Test

The Mathematics Continuous Performance Test (MATH–CPT) is a new measure designed in a different way from other CPT-type tests. Thus far, to the best knowledge of the authors, there is no other CPT-type test using mathematical problems as stimuli. There are three major differences between this new test and other CPT-type tests: (1) the stimuli shown require more complex cognitive processing of calculations, so responding demands more time. The intention of this change was to resolve floor effects found in many CPT tests in which stimuli are simple (Halperin, *et al.*, 1991; Greenberg, *et al.*, 1999; Pei-Chun, *et al.*, 2005); in these other tests, one unintentional mistake may cause the person tested to be diagnosed as having ADHD. (2) The second procedure used in the MATH–CPT is that in the analyses of the results, the test is divided into three parts of 150 problems each; each of these parts is then divided into three more parts of 50 problems each (for a total of nine parts per testee). Each subdivision of the nine parts includes measures of response time, consistency in response time, impulsive responses, and number of correct responses. This procedure allows better analysis of possible difficulties in sustained attention. (3) No time limit is set in the MATH–CPT. In other CPTs, the presentation of the stimuli appears on the screen for approximately 200 to 500 msec. Then an interstimulus interval is given, usually between 1 to 4 seconds, followed by the testee response, and then another stimulus is shown briefly. The testee is forced to respond within the designated inter-stimulus interval. In the MATH–CPT, the stimulus changes to the next only after the testee responds to the problem presented on the screen. The change to this self-paced procedure in the MATH–CPT is intended to encourage the natural tendency of the person tested to answer either quickly or slowly, versus forcing the testee to always respond quickly, as in other CPT-type tests. In 2001, Milich, Ballentine, and Lynam introduced the term slow cognitive tempo, which describes a behavioral pattern appearing in some children with ADHD, characterized by passivity or daydreaming or by hypoactive, confused, slow-moving, and sluggish responses. A self-paced procedure, as used in the MATH–CPT, was intended to assess slow cognitive tempo more accurately.

Barkley (2006) postulated the possibility of two distinct dimensions of inattention. One is described by the symptoms of the *Diagnostic and Statistical Manual of Mental Disorders* (APA, 2000) and can be characterized mostly by distractibility. The second dimension can be described as a passive and lethargic form and is consistent with the construct of slow cognitive tempo. It is argued that the self-paced procedure used in the MATH–CPT can distinguish better the second type of inattention. Having a forced procedure of limited time to respond, as in other CPT-type tests, is prompting or serving as a cue to the testee to be alert and respond quickly, often against a natural sluggishness, passivity, or daydreaming. In other words, other CPT tasks may, in fact, disturb the natural tendency of the testee, forcing attentiveness to the task at hand. It is hypothesized that the new MATH–CPT can better recognize this slow cognitive tempo tendency of people with this type of attention difficulties.

Purpose of the Study

The purpose of the present study is to develop and validate a new type of CPT by incorporating the three changes mentioned above. More specifically, the intent is to assess split-half and test-retest reliability and construct validity of this new measure to assess attention, impulsivity, reaction time, and sustained attention, in a sample of normal population and in a clinical sample of adolescents and adults with ADHD.

Method

Participants

Two hundred and forty participants (114 men, 126 women; *M* age = 18.9 yr., *SD* = 4.5) participated in the initial development of the MATH–CPT and served as a normal control group. The participants volunteered to participate in the study after they received an explanation of the nature of the study. All of them were used to assess split-half reliability. Out of this group, 39 participants (10 men, 29 women; *M* age 22.9 yr., *SD* = 5.3), were tested twice over the time period of one week to assess test-retest reliability. An additional 108 participants (55 men, 53 women; *M* age 21.6 yr., *SD* = 1.9) were administered the d2 Test of Attention after taking the MATH–CPT. Twenty-seven participants (20 men, 7 women; *M* age = 20.7 yr., *SD* = 5.4) took the WAIS–R (Wechsler, 1981; a more recent WAIS test was not standardized in Israel) and the Stroop Color and Word Test (Golden, 1978; Golden & Freshwater 1998). Among the 240 participants, there were 66 normal control participants (48 men, 18 women; *M* age 17.9 yr., *SD* = 5.0) who responded only to the MATH–CPT.

Sixty-three participants diagnosed as having ADHD were the clinical group in an exploratory validation process (36 men, 27 women; *M* age = 16.7 yr., *SD* = 4.8). These participants were recruited from various clinics and schools in northern Israel. The clinical group included only those who were diagnosed with ADHD according to the criteria of the APA (2000) by psychiatrists or clinical psychologists. They were diagnosed with one of the three diagnoses of ADHD: (1) ADHD combined type; (2) ADHD, predominantly inattention type; and (3) ADHD, predominantly hyperactive-impulsive type. In the present study, no statistics are available for the different sub-types of ADHD. Such classification is not accurate, and most of the time, it is not done in research using CPT-type tests. Most of the participants with ADHD were being treated with medication. Those participants agreed not to take their medication for 24 hr. prior to being tested by the MATH–CPT. Participants who had dyscalculia were excluded from the study based on their own report of having this disability. Of this sample, 37 of the participants with ADHD (23 women, 14 men; M age = 19.0 yr., SD = 5.7), in addition to being tested with the MATH–CPT, were tested with another CPT-type test, the Test of Variables of Attention (TOVA).

In this study, only participants older than 12 years were allowed to participate. A prior pilot study indicated that testing persons younger than 12 years could contribute inaccurate results, given the difficulty of the task or because the participants were unable to concentrate through the length of the test. This was true mainly for participants who had ADHD.

Materials

The MATH–CPT was programmed on a PC computer using the programming language Quickbasic. There are 450 simple mathematical problems of addition, subtraction, multiplication, and division, when the answer to the problem was not larger than nine. In the MATH–CPT test, the stimulus of a mathematical problem was shown on the computer screen with an answer (e.g., $1+3=4$ or $7-3=2$). The computer used a 17-in. screen with white numbers in an approximately 1.0-cm^2 matrix on a black background. Analysis included four main elements: (1) a measure of attention based on the number of correct answers; and (2) a measure of impulsivity based on the anticipated answers (or false alarm responses), i.e., a response was given in less than 500 msec. after the onset of the stimulus (a pilot study confirmed that it takes at least 500 msec. to process a simple mathematical problem effectively). The measure of impulsivity also included "fast wrong responses." They were wrong answers given faster than the participant's mean response time. (3) The measure of reaction time was based on the average time it took the participant to respond after the onset of the mathematical problem on the computer screen. (4) A measure of consistency in reaction time was assessed by the standard deviation of the reaction time.

To assess sustained attention, the scores on three one-thirds of the tests of 150 problems were used each time, each of these parts was then divided into three more parts of 50 problems each (for a total of nine parts per participant). There were 48 measures (scores on thirds and ninths of the test) for each one the four main elements of the MATH–CPT: measure of attention, measure of impulsivity, measure of reaction time, and measure of consistency in reaction time. These assessments allowed the tester to see the changes in sustained attention as the test progressed. This assessment of sustained attention is done by subjective observation of the responses to check for a decline, improvement, or no change in the personal measures along the progressive thirds or ninths of the test. This is a common procedure performed in many other CPT-type tests (Halperin, *et al.*, 1991; Conners, 2000).

All tests were administered before noon to minimize diurnal variations. Each session was performed in a distraction-free office while the tester was present. The test began with 20 practice problems administered after instructions were given orally. The wordings given to each subject are available from the authors.

For the purpose of exploratory validation, four different tests were selected for detection of attentional difficulties based on their prior construct validity (Lezak, 1995). The tests used for exploratory validation were the d2 Test of Attention, the Freedom from distractibility factor of the WAIS–R, the Stroop Color and Word Test, and the Test of Variables of Attention.

The d2 Test of Attention (Brickenkamp & Zillmer, 1998).—This is a cancellation test constructed by Brickenkamp in Germany in 1962 and validated again in 1998 by Brickenkamp and Zillmer. It has been used in many studies to assess attention. Qualities of the test include a test-retest reliability of above .90 in numerous studies. Validation has yielded statistically significant correlations with other measures of attention. Several studies have shown that a population with ADHD scores lower in various measures of the d2 Test of Attention (Semrud-Clikeman, Teeter, Parle, & Connor, 1995; Lufi, 2001). The test has the following variables: total items processed; total items marked minus mistakes; fluctuation rate, the longest row marked minus the shortest row marked (FR); formula calculating the percent of errors (total number of errors multiplied by 100, divided by the total number of items processed); and formula calculating concentration performance (CP: sum of correctly crossed-out items minus errors of the commission).

The freedom from distractibility factor of the WAIS–R was constructed by Kaufman (1975), who divided the sub-tests of the various Wechsler tests into three different factors rather than the original two main measures of Verbal and Performance IQs. Similar structures were reported for adult populations (Atkinson & Cyr, 1984; Kunce & De Vales, 1986; Alm & Kaufman, 2002). The freedom from distractibility factor has three subtests: Arithmetic, Coding, and Digit-span. This factor could identify attention and concentration difficulties in ADHD (Lufi & Cohen, 1985; Cohen, Becker, & Campbell, 1990; Anastopoulos, Spisto, & Maher, 1994) and in children with learning disabilities (Johnson & Blalock, 1987; Gregg, Hoy, & Gay, 1996).

The Stroop Color and Word Test (Golden, 1978; Golden & Freshwater, 1998).—This test is based on the original test by Stroop constructed in 1935. In this study, a Hebrew version of the test was used, as adapted by Lufi, Cohen, & Parish-Plass (1990). During the construction of the test, test-retest reliability values of .73 to .86 have been reported in various studies. In the past, Stroop Color–Word Test differentiates ADHD from a control group (Lufi, *et al.*, 1990; Grodzinsky & Diamond, 1992; Grodzinsky & Barkley, 1999).

The Test of Variables of Attention (TOVA; Greenberg, *et al.*, 1999).—This is a well-known commercial CPT-type test used to diagnose ADHD and to assess the efficacy of medications in children with ADHD. The TOVA uses geometric designs as stimuli. A special device is attached to the computer, and a child responds on a button located on the device. The stimulus is two geometric figures: a small white square which appears in the upper part of a larger black square in the target stimulus; the person tested has to quickly press the button. The non-target stimulus is the small white square which appears sometimes in the lower part of the larger black square, and the button should not be pressed. The stimulus appears for 100 msec., with inter-stimulus interval of 2 sec. The tests have 648 stimuli which last for 21.71 min. Before the start of the test, the person who is tested receives 3 min. of practice to be familiarized with the task. Norms from age 4 to 80 years were developed for the test. Also, the test reports the results in four quarters: (a) commission mistakes (those mistakes of the button pressed when the non-target stimulus appears), (b) omissions (those mistakes when the button is not pressed but the target stimulus appears), (c) response time, and (d) various measures of variability of response time. The most important score is an ADHD Score, which allows the tester to know if the testee is within the normal population or in the ADHD range. Test-retest reliability has indicated correlations of .77 to .99 for the various measures of the TOVA, while receiver operator characteristic (ROC) analysis showed sensitivity of 80% and specificity of 20% (Greenberg, *et al.*, 1999).

Procedure

After programming of the MATH–CPT, it was administered to a sample of a normal population. When performing the test, participants had to decide whether an answer was correct or incorrect. If the decision was that the answer was correct, then the participant pressed "1" on the numeric pad of the computer keyboard. If the testee decided that the answer was wrong, then he pressed "2." Another problem was shown only when an answer had been given to the previous problem. This sample was recruited from sources in northern Israel, including a local college, high schools, middle schools, and psychological treatment centers. Thirty-nine participants (10 men, 29 women; *M* age 22.9 yr., $SD = 5.3$) took the MATH–CPT again after one week to assess test-reliability. The d2 Test of Attention was administered to 108 participants of the normal control group immediately after they had taken the MATH–CPT. Twenty-seven completed the

WAIS–R test and the Stroop Color and Word Test in a randomized order either before or after the administration of the MATH–CPT. The individuals with ADHD were recruited from northern Israel high schools, colleges, and clinics to take the MATH–CPT as part of a special session during the morning hours.

The TOVA (Greenberg, *et al.*, 1999) and the MATH–CPT were administered to 37 adolescents and adults who were diagnosed with ADHD, based on the APA guidelines (2000), by a psychiatrist, neurologist or a psychologist. Half of the participants took the MATH–CPT first, while the other half completed the TOVA first. This procedure was used to counterbalance a possible practice effect. None of the participants used any medications at least 24 hours prior to completing those tests.

Results

Sample of Normal Population

The first step was to test a sample of a normal population with the new MATH–CPT and compare the mean scores obtained by the different sexes. Comparison of 116 men and 124 women showed no differences on the main measures of the MATH–CPT. Given these findings, the present authors could continue the development the MATH–CPT without regard to sex.

Assessing the effect of age on mean scores identified minor differences across ages 12 to 30 years in all measures of the MATH–CPT. One-way analysis of variance with age as a factor indicated no statistically significant differences among the four main measures of the MATH–CPT tested and on the overall attention measure developed using the data from this study (see Table 1).

Reliability

Split-half reliability of the MATH–CPT with responses by 240 normal testees showed the following estimates (r) for four of the main variables of the MATH–CPT: total time of the test ($r = .92$), standard deviation of reaction time ($r = .84$), correct answers ($r = .70$), and impulsivity ($r = .54$). Test-retest reliability was assessed with 39 students (28 women, 11 men; *M* age 22.9, $SD = 5.3$) who were tested twice, one week apart. The test-retest correlations of the main measures of the MATH–CPT were .85 for total time, .79 for standard deviation of reaction time, .78 for total correct responses, .75 for impulsive responses, and .48 for overall attention measure. Reliability of the measures of thirds and ninths were similar in magnitude to those from the main measures of the MATH–CPT.

Exploratory Validation

The first part of the exploratory validation was a comparison of 240 normal control participants with the 63 who were diagnosed with ADHD.

TABLE 1
MEANS AND STANDARD DEVIATIONS OF MAIN MEASURES OF MATH–CPT BY AGE ($N = 240$)

Age/Variable	*N*	Time		Correct Responses		Impulsivity		*SD* Time		Overall Attention Measures	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
12.00–12.99	8	845	273	430	10.39	12.20	6.30	1.00	0.69	−0.54	0.82
13.00–13.99	24	812	175	429	11.18	13.80	8.17	0.25	0.53	−0.70	0.53
14.00–14.99	13	776	156	431	7.57	14.22	6.80	0.69	0.33	−0.93	0.49
15.00–15.99	21	714	167	428	9.63	19.88	8.58	0.72	0.37	−0.78	0.66
16.00–16.99	23	767	208	431	8.18	12.50	6.94	0.88	0.50	−0.88	0.57
17.00–17.99	9	811	139	432	10.87	16.67	8.62	0.83	0.25	−0.83	0.44
18.00–18.99	11	740	191	428	12.65	14.80	7.21	0.98	0.63	−0.98	0.63
19.00–19.99	11	864	249	425	12.26	15.17	8.29	1.11	0.64	−0.49	0.94
20.00–20.99	17	749	188	431	10.00	14.32	6.95	0.82	0.39	−0.91	0.77
21.00–21.99	16	697	117	431	8.90	14.38	6.95	0.69	0.21	−0.95	0.48
22.00–22.99	16	719	142	429	13.15	15.43	9.29	0.74	0.28	−0.83	0.64
23.00–23.99	22	796	185	428	13.16	16.17	8.63	0.87	0.44	−0.52	0.88
24.00–24.99	19	801	218	435	9.30	12.35	5.93	0.87	0.49	−0.78	0.64
25.00–25.99	9	750	146	428	10.29	15.89	6.99	0.70	0.21	−0.89	0.70
26.00–30.00	21	737	156	434	7.53	11.63	5.10	0.73	0.29	−0.89	0.64
F		1.81		1.01		1.10		1.22		1.26	

$^{*}p < .05$.

A comparison of this target population of the MATH–CPT to normal control participants was set to develop overall attention measure aimed to diagnose ADHD based on the discriminant function equation. In this comparison, an alpha of .01 was used to reduce Type 1 error. There were statistically significant differences between groups on all the main measurements of the MATH–CPT (total time, *SD* time, total correct, impulsivity, and overall attention measure). Of the remaining 48 measures in the MATH–CPT (divisions of thirds and ninths of reaction time, standard deviations of reaction time, impulsive responses and number of correct responses), 38 showed statistically significant differences at .01. The comparisons of the main five measures of the MATH–CPT are shown in Table 2.

Stepwise discriminant function analysis was performed with 240 normal control participants compared with 63 individuals with ADHD to develop the overall attention measure, based on the discriminant function equation. Three of the main variables of the MATH–CPT [in order: total time (Wilks' lambda = .87), total correct responses (Wilks' lambda = .77), and total anticipatory responses (Wilks' lambda = .77)] correctly classified 91.6% of the participants in the two groups.

ROC was constructed to verify the cut-point for the final overall attention measure. This procedure assesses the sensitivity and specificity levels of the formula. Using a cut-point of 0 (set by the discriminant function analysis) gave an area under the curve (AUC) of .89, with high sensitivity of 88.5% (the probability that positive cases are correctly classified) and high specificity of 83.5% (the probability that negative cases are correctly classified). The asymptotic 95% confidence interval of the AUC was between .83 and .95.

To validate the new MATH–CPT, its scores were compared with an independent measure of attention, the d2 Test of Attention (Brickenkamp & Zillmer, 1998). This comparison yielded statistically significant correlations for three out of five main measures of the MATH–CPT: total time,

TABLE 2

MEANS, STANDARD DEVIATIONS, *F* RATIOS, EFFECT SIZES, AND POWER COMPARISON FOR NORMAL CONTROL GROUP ($n = 240$) TO THE ADHD GROUP ($n = 63$) MAIN MEASURES OF MATH–CPT

MATH–CPT	Normal Control Group		ADHD Group		t	Cohen's d
	M	*SD*	*M*	*SD*		
Total time	722.04	181.36	1,219.25	500.53	6.79*	0.96
SD time	0.82	0.42	1.92	1.53	5.51*	0.78
Total correct	429.73	10.41	414.98	21.47	5.15*	0.73
Impulsivity	14.40	7.60	24.21	13.65	3.74*	0.53
Overall attention measure	−0.78	0.67	1.46	1.58	10.68*	1.51

Note.—Cohen's d (1988) rule signals 0.80 and above as large effect. *$p < .001$.

standard deviation of reaction time, and overall attention measure, with five main measures of the d2 Test of Attention: total items processed; total items marked minus mistakes; fluctuation rate, the longest row marked minus the shortest row marked; formula calculating the percent of errors; and formula calculating concentration performance. The correlations ranged between −.51 and .53 (see Table 3).

The second exploratory validation was performed by correlating scores on the overall attention measure of the MATH–CPT with those on the freedom from distractibility factor (Kaufman, 1975) of the WAIS–R (Wechsler, 1981). The correlation was −.44 ($r = -.44$, $p < .05$; 95%$CI = -.34$, $-.54$). The third exploratory validation was performed by correlating scores on the Stroop Color–Word sub-test of the Stroop Color and Word Test with those on the overall attention measure of the MATH–CPT. This correlation was −.41 ($r = -.41$, $p < .05$; 95%$CI = -.31$, $-.51$). The fourth exploratory validation was a correlation of scores on the MATH–CPT with those on the TOVA for participants who were diagnosed as having ADHD; one statistically significant correlation indicated the possibility of having ADHD in both tests ($r = .39$, $p < .02$; 95%$CI = .32$, .46), but no other significant correlations were found for the other measures from both tests. For the present sample, MATH–CPT diagnosed correctly 22 of the 37 participants (59.5% of sensitivity), while the TOVA diagnosed correctly 21 of 37 participants (56.8% of sensitivity).

TABLE 3
CORRELATIONS FOR MAIN MEASURES OF MATH–CPT WITH MAIN MEASURES OF THE d2 TEST OF ATTENTION ($N = 108$)

MATH–CPT/d2 Test	Items Processed	Total Marked	FR	% Errors	CP
Total time	−.47†	−.51†	.39†	.29†	.53†
SD time	−.41†	−.45†	.43†	.28†	−.48†
Total correct	−.09	−.05	−.17	−.18	.01
Impulsivity	−.04	−.14	.31†	.30	.08
Overall attention measure	−.30†	−.35†	.38†	.29†	−.40†

*$p < .05$. †$p < .01$.

DISCUSSION

Steps taken to develop a new computerized tool to assess various aspects of attention and, specifically, to help in the diagnosis of ADHD were described. The study included different phases of assessing reliability (split-half and test-retest), validity, and creating a sample norm for a normal sample.

Validation of the MATH–CPT with the d2 Test of Attention (Brickenkamp & Zillmer, 1998) showed significant correlations with several measures of both tests. Although the correlations were statistically significant, they were regarded as modest. A possible explanation for these results was that the MATH–CPT, a computerized long test that relied on visual stimuli, probably tapped aspects of attention different from those of the d2 Test of Attention. The d2 Test of Attention was a much shorter test and relied on grapho-motor skills. It was possible that these two different cognitive tasks tapped different dimensions of attention, e.g., the distractibility type of inattention measured by the d2 Test of Attention, as compared to the sluggish cognitive tempo type of inattention (Barkley, 2006) measured by the MATH–CPT.

An important element of the exploratory validation was a comparison of the results of the MATH–CPT of the normal population group with those of the target population of ADHD. On all the measures of the MATH–CPT, the normal control group achieved better mean scores than the ADHD group. These findings allowed for the creation of the overall attention measure, a formula derived from the results to assess the overall performance on the test. The formula allowed correct classification of 91.6% of the participants. These percentages appeared to be satisfactory when one considers the report of other well-established CPTs. The TOVA test has reported classification accuracies of 84 and 86% for ADHD and control groups in two different studies (Greenberg, *et al.*, 1999). Using the Conners' CPT–II (Conners, 2000), correct classification of various groups was reportedly between 82 and 92%. Other studies with different CPT-type tests reported lower correct classifications in the classification of ADHD and non-ADHD populations (Matier-Sharma, Perachio, & Newcorn, 1995; Doyle, *et al.*, 2000). It may be noted that any single measure, including CPT-type tests, should not be used as a sole criterion for the diagnosis of ADHD. To improve accurate diagnosis of ADHD, a multi-modal approach to assessment should be adopted. More accurate diagnosis may be obtained by using additional diagnostic tools such as a review of personal histories, medical evaluations, and questionnaires administered to the parents, teachers, and the client.

The last exploratory validation of the study was the comparison of the MATH–CPT to the TOVA. The results showed a significant correlation between the two tests ($r = .39$), although they can be considered moderate correlations. Possible explanation for this result is as follows: although both are CPT-type tests, each one assesses attention in a different way. The TOVA has a graphic stimuli, has a limited time to answer, and lasts longer. In contrast, the MATH–CPT has an open time format, and the stimuli of simple mathematical problems is much harder. It is important to remember the fact that in the present study, the MATH–CPT was able to identify the ADHD sample somewhat better as compared to the TOVA (59.5% for the MATH–CPT, 56.8% for the TOVA).

Clinicians using the MATH–CPT should consider possible errors in the test, false negatives (diagnosing individuals with ADHD as non-ADHD), and false positives (incorrectly concluding that a non-client child has ADHD). This issue was clarified by the results of the discriminant function analysis. In this analysis, 91.6% of the participants were classified correctly. Within this percentage, the sensitivity (the probability that positive cases are correctly classified), i.e., the rate at which the diagnostic procedure detects a disorder when it is present, was 78.3% for the ADHD group. In the normal control group, the specificity (the ability of the test to detect the absence of a disorder when it was not present) was 94.3%.

The percentages described above were achieved when the value of the overall attention measure was set at 0. Since false negative and false positive vary inversely, one can increase or decrease the false negative and false positive percentages by changing the cutting score for the measure of attention. A more positive number (>0) could reduce the possibility of false positives, while a more negative number (<0) could reduce the possibility of false negatives. With this information, the clinician can reach the proper decision about a diagnosis of ADHD. Based on the results presented, it is suggested that the score of 0 should be used as the cutoff line for diagnosis. When a number is <0, it is in the non-ADHD range. Conversely, when the number is >0, it falls in the ADHD range. However, any raw score between −0.3 and 0.3 (about one-half the standard deviation) should be considered as borderline for diagnosis of ADHD. A raw score >0.3 can more certainly result in the diagnosis of ADHD, while a raw score of <-0.3 is more likely not to be diagnosed as non-ADHD. In the present study, discriminant function analysis assessed this issue and was confirmed by ROC statistics.

Users of the MATH–CPT should be aware of one more issue resulting from data mentioned here. The test-retest correlation for the overall measure of attention ($r = .48$), although significant, can be considered low. One possible explanation for this finding is that a one-time classification of the testee's score might have sufficient diagnostic power. However, a repetition of the MATH–CPT by the same person, within a short period, could decrease the specificity and sensitivity of the results.

A close examination of Table 2 reveals that the standard deviations of all the variables used in the study were much larger in the ADHD group as compared to the normal control population. These findings reflected the nature of the difficulties of the ADHD group, which is inconsistency in the functioning of many areas of life. Greenberg, *et al*. (1999), in their study of the TOVA test, stated that 80% of the variance in their study was explained by the variability of reaction time.

Limitations of the Study

Because the task in the MATH–CPT is based on a learned skill, i.e., being able to calculate, exploring the relation between mathematical skills and ability to concentrate becomes focal. In other studies, children with ADHD solved mathematical problems more slowly and made more mistakes than a control group (Zentall, 1990; Zentall & Smith, 1993; Zentall, Smith, Lee, & Wieczorek, 1994). Such differences were found even when the groups were matched for IQ, reading comprehension, and motor ability. This possible difficulty in interpretation of the MATH–CPT must be investigated. Until data are available, users of the MATH–CPT should be cautious in diagnosing individuals with dyscalculia as having ADHD.

Norm groups were not sufficiently established for every age group, something that requires attention. Future studies with the MATH–CPT should include attention to the severity of ADHD and to the different subtypes of ADHD. The use of the new instrument with different clinical groups should be investigated in the future, as well as the estimations of inter-relations among measures of attention, impulsivity, reaction time, and sustained attention.

REFERENCES

Alm, J., & Kaufman, A. S. (2002) The Swedish WAIS–R factor structure and cognitive profiles for adults with dyslexia. *Journal of Learning Disabilities,* 35, 321-334.

Aman, M., Kern, R., McGhee, D., & Arnold, E. (1993) Fenfluramine and methylphenidate in children with mental retardation and Attention Deficit Hyperactivity Disorder: laboratory effects. *Journal of Autism and Developmental Disorders,* 23, 491-506.

American Psychiatric Association. (2000) *Diagnostic and statistical manual of mental disorders.* (4th ed., text rev.) Washington, DC: Author.

Anastopoulos, A. D., Spisto, M. A., & Maher, M. C. (1994) The WISC–III freedom from distractibility factor: its utility in identifying children with Attention Deficit Hyperactivity Disorder. *Psychological Assessment*, 6, 368-371.

Atkinson, L., & Cyr, J. (1984) Factor analyses of the WAIS–R: psychiatric and standardization samples. *Journal of Consulting and Clinical Psychology,* 52, 714-716.

Barkley, A. R. (2006) *Attention-Deficit Hyperactivity Disorder: a handbook for diagnosis and treatment.* New York: Guilford.

Bennett, D. E., Zentall, S. S., & French, B. F. (2006) The effects of computer administered choice on students with and without characteristics of Attention Deficit/Hyperactivity Disorder. *Behavioral Disorders,* 31, 189-203.

Borgaro, S., Pogge, D. L., DeLuca, V. A., Bilginer, L., Stokes, J., & Harvey, P. D. (2003) Convergence of different versions of the Continuous Performance Test: clinical and scientific implications. *Journal of Clinical and Experimental Neuropsychology,* 25, 283-292.

Brickenkamp, R., & Zillmer, E. (1998) *The d2 Test of Attention.* Seattle, WA: Hogrefe & Huber.

Cohen, J. (1988) *Statistical power analysis for the behavioral sciences.* New York: Wiley.

Cohen, M., Becker, M. G., & Campbell, R. (1990) Relationships among four measures of assessment of children with attention deficit-hyperactivity disorder. *Journal of School Psychology,* 28, 189-202.

Conners, C. K. (2000) *Conners' Continuous Performance Test–II: technical guide and software manual*. North Tonawanda, NY: Multi-Health Systems.

Cronblatt, B. A., Risch, N. J. Faris, G., Friedman, D., & Erlenmeyer-Kimling, L. (1988) The Continuous Performance Test, identical-pairs version (CPT–IP): I. New findings about sustained attention in normal families. *Psychiatry Research,* 26, 223-238.

Doyle, A. E., Biederman, J., Seidman, L. J., Weber, W., & Faraone, S. (2000) Diagnostic efficiency of neuropsychological test scores for discriminating boys with and without Attention Deficit-Hyperactivity Disorder. *Journal of Consulting and Clinical Psychology,* 68, 477-488.

Golden, C. J. (1978) *The Stroop Color and Word Test: a manual for clinical and experimental uses.* Chicago: Stoelting.

Golden, C. J., & Freshwater, S. M. (1998) *Stroop Color and Word Test, Revised: adult manual for clinical and experimental uses.* Wood Dale, IL: Stoelting.

Gordon, M. (1987) *The Gordon diagnostic system.* DeWitt, NY: Gordon Systems.

Greenberg, L. M., Kindschi, C. L., & Corman, C. M. (1999) *T.O.V.A. Test of Variables of Attention: clinical guide.* Los Alamitos, CA: Universal Attention Disorders.

Gregg, N., Hoy, C., & Gay, A. F. (1996) *Adults with learning disabilities: theoretical and practical perspectives.* New York: Guilford.

Grodzinsky, G. M., & Barkley, R. A. (1999) The predictive power of frontal lobe tests in the diagnosis of Attention Deficit Hyperactivity Disorder. *Clinical Neuropsychologist,* 13, 12-21.

Grodzinsky, G. M., & Diamond, R. (1992) Frontal lobe functioning in boys with Attention-Deficit Hyperactivity Disorder. *Developmental Neuropsychology,* 8, 427-445.

Halperin, M. H., Sharma, V., Greenblatt, E., & Schwartz, S. (1991) Assessment of the Continuous Performance Test: reliability and validity in a nonreferred sample. *Psychological Assessment: a Journal of Consulting and Clinical Psychology,* 3, 603-608.

Johnson, D. J., & Blalock, J. W. (1987) Summary of problems and needs. In D. J. Johnson & J. W. Blalock (Eds.), *Adults with learning disabilities: clinical studies.* Orlando, FL: Grune & Stratton. Pp. 277-296.

Kaufman, A. S. (1975) Factor analysis of the WISC–R at 11 ages between $6^1/_2$ and $16^1/_2$ years. *Journal of Consulting and Clinical Psychology,* 48, 135-147.

Kunce, J. T., & De Vales, E. S. (1986) Cross-cultural factor analytic similarity of Wechsler Intelligence scores for Mexican adults. *Journal of Clinical Psychology,* 42, 165-169.

Lezak, M. D. (1995) *Neuropsychological assessment.* (3rd ed.) New York: Oxford Univer. Press.

Lufi, D. (2001) Double-coding Test: a new paper-and-pencil measure of eye-hand coordination. *Perceptual and Motor Skills*, 92, 815-826.

Lufi, D., & Cohen, A. (1985) Using the WISC–R to identify Attention Deficit Disorder. *Psychology in the Schools,* 22, 40-42.

Lufi, D., Cohen, A., & Parish-Plass, J. (1990) Identifying Attention Deficit Hyperactive Disorder with the WISC–R and the Stroop Color and Word Test. *Psychology in the Schools,* 27, 28-34.

Mahone, E. M. (2005) Measurement of attention and related functions in the preschool child. *Mental Retardation and Developmental Disabilities Research Reviews,* 11, 216-225.

Matier-Sharma, K., Perachio, N., & Newcorn, J. H. (1995) Differential diagnosis of ADHD: are objective measures of attention, impulsivity, and activity level helpful? *Child Neuropsychology,* 1, 118-127.

Milich, R., Ballentine, A. C., & Lynam, D. R. (2001) ADHD/combined type and ADHD/predominantly inattentive type of distinct and unrelated disorders. *Clinical Psychology: Science and Practice,* 8, 463-488.

Nuechterlein, K. H., Edell, W. S., Norris, M., & Dawson, M. E. (1986) Attentional vulnerability indicators, thought disorder, and negative symptoms. *Schizophrenia Bulletin*, 12, 408-426.

Pei-Chun, H., Ching, L. C., Yen, K. Y., Yi, C. Y., Tzung, L. Y., I-Hui, L., & Po, S. C. (2005) Norms of performance of sustained attention among a community sample: Continuous Performance Test study. *Psychiatry and Clinical Neurosciences,* 59, 170-176.

Riccio, C. A., Reynolds, C. R., & Lowe P. A. (2001) *Clinical applications of Continuous Performance Tests: measuring attention and impulsive responding in children and adults.* New York: Wiley.

Semrud-Clikeman, M., Teeter, P. A., Parle, N., & Connor, R. T. (1995) Innovative approach for working with children with ADHD. Paper presented at the annual conference of the American Educational Research Association, San Francisco, CA, April.

Soreni, N., Crosbie, J., Ickowicz, A., & Schachar, R. (2009) Stop signal and Conners' continuous performance tasks: test-retest reliability of two inhibition measures in ADHD children. *Journal of Attention Disorders*, 13, 137-143.

Wechsler, D. (1981) *Wechsler Adult Intelligence Scale–Revised manual.* New York: The Psychological Corporation.

Zentall, S. S. (1990) Fact retrieval automatization and math problem solving: learning disabled, attention disordered, and normal adolescents. *Journal of Educational Psychology,* 82, 856-865.

Zentall, S. S., & Smith, Y. N. (1993) Mathematical performance and behavior of children with hyperactivity with and without coexisting aggression. *Behavior Research and Therapy,* 31, 701-710.

Zentall, S. S., Smith, Y. N., Lee, Y. B., & Wieczorek, C. (1994) Mathematical outcomes of attention deficit hyperactivity disorder. *Journal of Learning Disabilities,* 27, 510-519.

Accepted December 9, 2011.