Usefulness and Validity of Continuous Performance Tests in the Diagnosis of Attention-Deficit Hyperactivity Disorder Children

Itai Berger1,*, Ortal Slobodin1, Hanoch Cassuto2

1Pediatric Neurology Unit, The Neuro-Cognitive Center, Hadassah Hebrew University Medical Center, Jerusalem, Israel
2Pediatric Neurology, Leumit and Clalit HMO, Jerusalem, Israel

*Corresponding author at: Pediatric Neurology Unit, The Neuro-Cognitive Center, Hadassah-Hebrew University Medical Center, Mount Scopus Campus, POB 24035, Jerusalem 91240, Israel. Tel.: 972-2-5844903; fax: 972-2-5328963.
E-mail address: itberg@hadassah.org.il (I. Berger).

Abstract

Objective: Despite the popularity of continuous performance tests (CPT) in supporting the diagnostic procedure of attention-deficit hyperactivity disorder (ADHD), these measures are still controversial mainly due to limited sensitivity, specificity, and ecological validity. Thus, there continues to be a need for further validation of these objective attention measures. The purpose of this study was to evaluate the usefulness of a CPT that includes environmental distracting stimuli, in supporting the diagnosis of ADHD in children.

Method: Participants were 798 children aged 7–12 years (493 boys and 305 girls). The ADHD group included 339 children, whereas the control group included 459 children without ADHD. The study employed the MOXO-CPT, which incorporates visual and auditory stimuli serving as environmental distractors.

Results: Compared to their unaffected peers, children with ADHD received significantly lower scores in all 4 CPT indices: attention, timing, hyperactivity, and impulsivity. Specifically, ADHD children were less attended to the stimuli and performed fewer reactions on accurate timing. Furthermore, children with ADHD performed significantly more impulsive and hyperactive responses than controls. Receiver operating characteristic analysis revealed fair to excellent diagnostic ability of all CPT indices except impulsivity, which showed poor ability to distinguish ADHD children from controls. The test’s total score yielded excellent diagnostic performance.

Conclusions: MOXO-CPT consistently distinguished between children with ADHD and their unaffected peers, so that children with ADHD performed worse than controls in all study indices. Integration of CPT indices improves the diagnostic capacity of ADHD and may better reflect the complexity and heterogeneity of ADHD.

Keywords: ADHD; Assessment; Attention; Childhood neurologic disorders; Test construction

Introduction

Attention-deficit hyperactivity disorder (ADHD) is an early onset, enduring neuro-developmental deficit, which is among the most prevalent chronic health conditions (Polanczyk, de Lima, Horta, Biederman, Rohde, 2007). Given the wide heterogeneity of symptoms, complex manifestation, and high comorbidity rates, there is an importance in a developmental perspective that views ADHD as a multifactorial disorder with multiple, causal processes and pathways (Berger, Remington, Leitner, & Leviton, 2015).

The diagnosis of ADHD combines clinical observation with subjective reports. In clinical practice, the assessment process in children typically involves obtaining information from multiple sources, including parents–teachers’ interviews or questionnaires, clinical assessment, and, when needed, neuropsychological or other testing (American Academy of Pediatrics [AAP], 2001; Seixas, Weiss, & Muller, 2012; Wolraich et al., 2011). The vulnerability of the clinical assessment, interviews, and questionnaires to clinician and informant biases (Rousseau, Measham, & Bathiche-Suidan, 2008; Serra-Pinheiro, Mattos, & Regalla, 2008; Skounti, Philalithis, & Galanakis, 2007) may lead to underdiagnosis as well as to overdiagnosis...
(and consequently to under or over treatment) of ADHD, particularly in certain groups (e.g., girls) (Ginsberg, Quintero, Anand, Casillas, & Upadhyayam, 2014; Quinn & Madhoo, 2014). The “transparency” of ADHD, even among experienced professionals, may cause misdiagnosis, harm the treatment aspects, and has potential non-beneficial prognostic aspects (Berger & Maier, 2014). Thus, there has long been interest in developing objective laboratory-based measures that could support clinical diagnosis. Currently, the most popular objective laboratory-based measures that could support clinical diagnosis are the continuous performance test (CPT) tasks (Edwards et al., 2007; Vogt & Williams, 2011). The majority of CPT tasks (e.g., CPT-Degraded Stimulus, Adler et al., 2001; Gordon Diagnostic System, Gordon & Mettelman, 1987; Test of Variables of Attention, Greenberg, 1997; CPT identical-pairs, Cornblatt, Risch, Faris, Friedman, & Erlenmeyer-Kimling, 1988) measure selective attention, sustained attention, and impulsivity, which can be used alongside clinical evaluation to inform the diagnostic process (Hall et al., 2014, 2016; Sonuga-Barke, Sergeant, Nigg & Willcutt, 2008). Usually, the test involves the rapid presentation of a series of visual or auditory stimuli over a period of time (typically numbers, letters, number/letter sequences or geometric figures). Subjects are instructed to respond to the “target” stimulus and to avoid responding to “non-target” stimuli. Response to non-target stimuli is referred to as a “commission error,” whereas the absence of response to target stimuli is referred to as an “omission error.” Omission errors have been empirically assumed to measure inattention, whereas commission errors are considered to measure impulsivity. Other standard measures of CPT responses include the number of correct responses, response time (RT), and the variability in RT.

Like other cognitive tests, the CPT may serve as an aiding tool in the diagnostic procedure but is still considered controversial by some authors mainly due to questions regarding limited sensitivity, specificity, and ecological validity (Arble, Kuentzel, & Barnett, 2014; Edwards et al., 2007; Nigg, Willcutt, Doyle, & Sonuga-Barke, 2005). The validity of a diagnostic test refers to what the test measures and how well it does so. Barkley (1991) observed that although CPTs present excellent face validity (measure the ability to sustain attention for an extended period of time), their psychometric validity are often low to moderate. However, psychometric properties of CPT vary considerably between different tests. An analysis of eight CPT studies revealed a wide heterogeneity in measures of sensitivity (9%–88%) and specificity (23%–100%) to ADHD. Notably, the studies with higher specificities (100% and 94%) had low sensitivities (13% and 62%, respectively) (Pan, Ma & Dai., 2007).

The controversial psychometric qualities of various CPT tasks were often attributed to a low ecological validity, namely the ability to simulate the difficulties of ADHD patients in everyday life (Neguț, Jurma, & David, 2016; Pelham et al., 2011; Rapport, Chung, Shore, Denney, & Isaacs, 2000). In contrast to many practitioners’ view that CPT provides a snapshot of child’s typical behavior, most CPTs are actually free from external, environmental distracting stimuli, which are thought to significantly impair the cognitive performance of ADHD children (APA, 2000). Although non-target stimuli can be often considered as distractors, they are actually an integral part of the go/no-go paradigm and could not and should not be filtered. In contrast, everyday environment is characterized by an ongoing flow of external non-relevant stimuli that have to be ignored or filtered in order to allow adaptive functioning. This lack of external stimuli in most CPTs is particularly crucial for older children, where prediction to school behavior may be desired. Thus, several authors emphasized the need to improve the ecological validity of the CPT by evaluating the child’s behavior in more natural settings (Barkley 1991; Neguț et al., 2016; Parsons, Bowerly, Buckwalter, & Rizzo, 2007).

The degree to which CPT corresponds to ADHD symptomatology remains unclear. For instance, the correlation between CPT performance and parent or teacher rating scales is low to moderate (Forbes, 1998; Rielly, Cunningham, Richards, Elbard, & Mahoney, 1999). It is possible that CPT is sensitive to only part of the core deficits of ADHD, such as inattention and impulsivity but not to others, such as hyperactivity. At the same time, CPT probably assesses unique aspects of ADHD, which are not captured by rating scales (Hall et al., 2016).

It has also been shown that developmental changes in ADHD symptomatology were not consistently reflected in changes in CPT performance (Halperin, Trampush, Miller, Marks, & Newcorn, 2008; Vaughn et al. 2011). One of the explanations for this inconsistency considers ADHD as a result of various neurological dysfunctions with different developmental trajectories. The first is a subcortical dysfunction that persists across development and the second is a frontal dysfunction that improves with time (Halperin & Schulz, 2006). CPT outcomes that are related to frontal lobe functionality (e.g., commission errors) should decrease over time concomitantly with ADHD symptoms, whereas CPT outcomes linked to subcortical functioning (e.g., RT variability) should remain relatively constant and unrelated to ADHD symptomatology.

Another limitation of many current versions of CPT is ceiling effect (Berlin, Bohlin, Nyberg, & Janols, 2004; Mahone, Pillion, & Hiemenz, 2001), which seems to be related to the ease of visual format CPTs (Lasee & Choi, 2013). Most CPT paradigms rely on cognitive stimuli that involve clearly distinctive target/non-target stimuli (e.g., letters or numbers). This easy distinctiveness often lead to high hit rates and low false alarm rates, or even perfect performance (See, Howe, Warm, & Dember, 1995). Ceiling effect due to insufficient cognitive demands is even more pronounced in older children and adolescents (Epstein et al., 2003). Despite this, most studies constrain their analyses to error rates even in the presence of highly accurate performance (Huang-Pollock, Karalunas, Tam, & Moore, 2012). Finally, there are limitations in CPT ability to
measure deficits in timing, which are often evidenced in ADHD (Castellanos & Tannock, 2002; Kuntsi, Oosterlaan, & Stevenson, 2001). Traditionally, difficulties in timing at a CPT task are evaluated by mean RT for correct responses to the target (which is interpreted as a measure of information processing and motor response speed) and by the standard deviation of RT for correct responses to the target (which is interpreted as a measure of variability or consistency) (Greenberg & Waldman, 1993). Previous findings revealed inconsistent results regarding RT in children with ADHD. Although some studies reported slower RT in ADHD than in healthy controls (Andreou et al., 2007; de Zeeuw et al., 2008), others found faster RT in ADHD children. Slow responses were attributed to lapses of attention (Hervey et al., 2006). Such results suggest that RT indices reflect multiple interacting processes, including attention, speed–accuracy trade-off effects, rate of information processing, motor bias, and response criteria (Ratcliff, 2002). Accurate assessment of these multiple interacting processes and understanding of how they contribute to an ultimate RT score are therefore critical to the interpretability of RT data (Huang-Pollock et al., 2012).

Taken together, there continues to be a need for an objective measure of attention that addresses some of the known shortcomings. The purpose of this study was to evaluate the usefulness of the MOXO-CPT as a supporting measure in the diagnosis of ADHD in children aged 7–12 years. Specifically, the study explores the discriminative validity of the test between children with ADHD and their non-affected peers. The term MOXO is derived from the world of Japanese martial arts and means a “moment of lucidity.”

With the purpose of improving ecological validity, the MOXO-CPT incorporates external interfering stimuli (auditory and visual) that serve as distractors. Previous research has shown that incorporating environmental distractors improved the test’s utility in differentiating ADHD from control children (Cassuto, Ben-Simon, & Berger, 2013). Importantly, the negative effect of environmental distractors on CPT performance of ADHD children did not diminish with age, suggesting that distractibility is a persistent deficit in ADHD (Slobodin, Cassuto, & Berger, 2015). Finally, the timing index of the MOXO-CPT is able to differentiate between motor speed problems and inattention.

Materials and Methods

Participants

Participants in this study were 798 children aged 7–12 years, of them 493 were boys and 305 were girls. The study group included 339 children diagnosed with ADHD (mean age (in years), 9.27, SD = 1.65), and the control group included 459 children without ADHD (mean age (in years) 9.71, SD = 1.64). Children were divided into six different age categories. The category of “8 years”, for example, included children who were 8 years or older but younger than 9 years. As can be seen in Table 1, within each age category, the two study groups did not differ in gender distributions.

Participants in the ADHD group were recruited from children referred to the outpatient pediatric clinics of a neurocognitive center, based in a tertiary care university hospital. The children were referred through their pediatrician, general practitioner, teacher, psychologist, or directly by the parents.

Inclusion criteria for participants in the ADHD group were: (a) The child met the criteria for ADHD according to DSM-IV-TR (APA, 2000), as assessed by a certified pediatric neurologist. The diagnostic procedure included a semistructured interview with the child and parents, completion of questionnaires, and medical/neurological examination that confirmed ADHD diagnosis; (b) the child scored above the standard clinical cutoffs for ADHD symptoms on ADHD/DSM-IV scales (DuPaul, Power, Anastopouls, & Reid, 1998); and (c) the child was drug-naïve.

Participants in the control group were randomly recruited from typically developed children who study in regular classes at primary schools. As a common practice in other similar studies (Advokat, Martino, Hill, & Gouvier et al., 2007; Cordier, Muaro, Wilkes-Gillan, Speyer, & Pearce, 2014) and for the purpose of this article (evaluate the usefulness of a CPT), a typically developing child was defined as a child who did not have childhood developmental disorder (i.e., scored below the clinical cutoff for any of the Conners’ Parent Rating Scales-Revised [CPRS-R] subscales and DSM-IV scales) and for whom no concerns had been raised about development by a teacher or health professional (Berlin et al., 2004; de Zeeuw et al., 2008). Inclusion criteria for participants in the control group were: (a) the child scored below the clinical cutoff point for ADHD symptoms on ADHD/DSM-IV scales. (b) the absence of academic or behavioral problems. For this purpose, the parents and teachers were asked to state (with the informed consent) that there are no behavioral/academic/social difficulties. Using parent and teacher ratings of ADHD symptoms is considered a common and reliable method for assigning participants to ADHD/non-ADHD groups (e.g., Uno et al., 2006; Van Mourik et al., 2007).

The control group of the study included only typically developed children and not referred children without ADHD. The reason for this selection is that referred children who are diagnosed as non-ADHD in clinical settings are frequently diagnosed
with other comorbid conditions. Because the study focused on the ability of the MOXO-CPT to differentiate ADHD children from their non-affected peers, there was a rationale for preferring a control group who was rigorously screened for ADHD. Nevertheless, the limitations of this selection will be further discussed.

The exclusion criteria for all participants were: intellectual disability, other chronic condition, chronic use of medications, and primary psychiatric diagnosis (e.g., depression, anxiety, and psychosis). All participants in this study studied in regular classes in regular schools.

Based on a community health center, our sample was multicultural and heterogeneous with regard to potentially confounding factors correlated with the diagnosis of ADHD (Wolraich et al., 2011). All participants agreed to participate in the study, and their parents provided a written informed consent to the study, approved by the Helsinki committee (IRB) of Hadassah-Hebrew University Medical Center, Jerusalem, Israel.

### Measurements

**Measurement of child behavior**—The parent and teacher forms of the *Conners Rating Scale* (Conners, 1997a, 1997b) or ADHD/DSM-IV Scales (DuPaul et al., 1998) were employed. These are of well-accepted validity and reliability, and both are regarded as standards in ADHD diagnosis (Barkley, 2014).

**MOXO-CPT**—This study employed the MOXO-CPT version (Berger & Goldzweig, 2010). The MOXO-CPT (NeuroTech Solutions Ltd) is a standardized computerized test designed to diagnose ADHD-related symptoms. The test consists of eight stages (levels). Each level consists of 53 trials and lasts 114.15s. The total duration of the test is 15.2 min. In each trial, a stimulus (target or non-target) is presented in the middle of the computer screen for durations of 0.5, 1, or 3s and is followed by a “void” of the same duration (Fig. 1). Fifty-three stimuli are presented in each level, of which 33 are target stimuli and 20 are non-target. By implanting a void period after each stimulus and using variable presentation durations of the elements, the MOXO-CPT could distinguish accurate responses performed in “good timing” (quick and correct responses to the target performed during stimulus presentation) from accurate but slow responses (correct responses to the target performed after the stimulus presentation, during the void period). These two aspects of timing correspond to the two different problems of ADHD, responding quickly and accurately (National Institute of Mental Health, 2012).

Each stimulus remains on the screen for the full duration of the designated presentation time, regardless of whether a response was given. This practice allows the measuring of the timing of the response as well as its accuracy. The screen size

### Table 1. Participants’ gender distribution

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>ADHD (N = 339)</th>
<th>Control (N = 459)</th>
<th>Difference χ² (1, N = 798)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td></td>
<td></td>
<td>1.14</td>
</tr>
<tr>
<td>Male</td>
<td>62 (64.52)</td>
<td>64 (57.81)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>22 (35.48)</td>
<td>30 (42.19)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td>0.37</td>
</tr>
<tr>
<td>Male</td>
<td>47 (70.15)</td>
<td>50 (66.67)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>20 (29.85)</td>
<td>25 (33.33)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td>0.0004</td>
</tr>
<tr>
<td>Male</td>
<td>37 (64.91)</td>
<td>46 (64.69)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>30 (35.09)</td>
<td>25 (35.31)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td>0.15</td>
</tr>
<tr>
<td>Male</td>
<td>35 (56.45)</td>
<td>47 (54.02)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>27 (44.55)</td>
<td>40 (45.98)</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>Male</td>
<td>32 (62.74)</td>
<td>55 (63.21)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>19 (37.26)</td>
<td>32 (27.79)</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>Male</td>
<td>24 (60)</td>
<td>44 (59.46)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>16 (40)</td>
<td>30 (40.54)</td>
<td></td>
</tr>
</tbody>
</table>

ADHD, attention-deficit hyperactivity disorder; N = number. *p < .05 **p < .01 ***p < .001.
is 125 high and 166 wide. The child is located 60 cm from the screen and is instructed to respond to target stimulus as quickly as possible by pressing the space bar once and only once. The child is also instructed not to respond to any other stimuli but the target and not to press any other key but the space bar.

**Test stimuli.** Both target and non-target stimuli are cartoon pictures free of letters or numbers. The absence of letters and numbers in the stimuli is important given the fact that ADHD children tend to have learning difficulties (e.g., dyslexia and dyscalculia), which may be confounded with CPT performance (Seidman, Biederman, Monuteaux, Doyle, & Faraone, 2001). Target stimulus is always a cartoon image of a child’s face. Non-target stimuli include five different images of animals. Both target and non-target stimuli are 41 mm × 41 mm large and are always presented in the center of the screen (Fig. 2).

**Distracting stimuli.** To simulate everyday environment, the MOXO-CPT includes visual and auditory distracting stimuli that are not part of the non-target stimuli. The distracting stimuli are of various degrees of similarity to the target stimulus. Distractors are short animated video clips containing visual and auditory features, which can appear separately or together.
All distractors are typical elements in the child environment, a unique feature to the MOXO-CPT. Overall, six different distractors are included, each of them could appear as pure visual (e.g., three birds moving their wings), pure auditory (e.g., birds singing), or as a combination of them (birds moving their wings and singing simultaneously). Each distractor was presented for a different duration ranging from 3.5s to 14.8s, with a fixed interval of 0.5s between two distractors. Visual distractors include six different stimuli: a gong (presented for 6.8s), a bowling ball (3.5s), birds (9.25s), a warrior (Jedi) with saber (14.8s), a saber (6.8s), and a flying airplane (8.6s).

Visual distractors appear at one of the four spatial locations on the sides of the screen: down, up, left, or right. Visual distractors that appear on the left or right axis are 200–400 pixels high and 100–200 pixels wide, whereas those that appear on the up or down axis are 100–200 pixels high and 100–600 wide. The distance between visual distractors and target or non-target stimuli is always 21 mm (Fig. 3).

Auditory distractors include the six corresponding sounds of all visual distractors (e.g., a gong sound, sound of a bowling ball, birds singing, etc.). The sound is delivered through loudspeakers located on both sides of the screen (about 60 cm distance from the child’s ears). The sound intensity is about 70% of the maximal intensity of the loudspeakers.

Distractors’ onset is not synchronized with target or non-target’s onset and could be generated during target or non-target stimulus or during the void period.

Test levels. The test consists of eight levels with 53 trials in each level. The stimuli and their presentation time are identical across all levels; however, the levels differ in the visual and auditory distractors present in the trials. Different levels of the MOXO-CPT are characterized by a different set of distractors: Levels 1 and 8 did not include any distractors but only target and non-target stimuli, Levels 2 and 3 contained pure visual stimuli, Levels 4 and 5 contained pure auditory stimuli, and Levels 6 and 7 contained a combination of visual and auditory stimuli. The sequence of distractors and their exact position on the display are constant for each level. The load of the distracting stimuli increases in the odd number levels. During the second, fourth, and sixth levels, only one distractor is presented at a time. During the third, fifth, and seventh levels, two distractors are presented simultaneously.

Performance indices. The MOXO-CPT includes four performance indices:

Attention The number of correct responses (pressing the key in response to a target stimulus), which were performed either during the stimulus presentation on the screen or during the void period that followed. Thus, it was possible to evaluate whether the participant responded correctly to the target (was attentive to the target) independently of how fast he was. Knowing how many responses are expected, it was also possible to calculate the number of times the target was presented, but the patient did not respond to it (omission errors).

The difference between the total number of the target stimuli and the number of correct responses produces the number of omission errors. The score in this index was calculated as the average of correct responses throughout the eight test levels.

Timing The number of correct responses (pressing the key in response to a target stimulus), which are given while the target stimulus is still presented on the screen. This index did not include responses that were performed during the void period (after the stimulus has disappeared). In contrast to the attention index that measures correct responses independently of their speed, the

![Fig. 3. MOXO-CPT visual distractors.](http://acn.oxfordjournals.org/ at Hebrew University, Jerusalem)
current index measures correct responses only if they were performed on good timing. The score in this index was calculated as the average of correct responses, while the target stimulus is still presented on the screen throughout the eight test levels.

**Impulsivity**  The number of commission errors performed only during the time in which a non-target stimulus is present on the screen. Usually, commission errors are coded in any case of inappropriate response to the target (e.g., pressing a random key) (Greenberg & Waldman, 1993). In contrast, the MOXO-CPT’s impulsivity index considers only the pressings on the keyboard’s space bar in response to non-target stimulus as impulsive behavior. All other non-inhibited responses (e.g., pressing the keyboard more than once) were not coded as impulsive responses (as will be described in Hyperactivity). Score in this index was calculated as the average of impulsive responses throughout the eight test levels.

**Hyperactivity**  The number of all types of commission responses that are not coded as impulsive responses (e.g., multiple responses, random key pressing). Several examples are: (a) Multiple responses—pressing the keyboard’s space bar more than once (in response to target/ non-target), which is commonly interpreted as a measure of motor hyperresponsivity (Greenberg & Waldman, 1993). The MOXO-CPT considered only the second press and above (the first response would be considered as correct response with good timing as multiple responses, as correct response with poor timing, or as impulsive response, depends on the type of element appearing on the screen). (b) Random key pressing—pressing any keyboard button other than the space bar. Multiple and random key pressing was chosen as a measurement of hyperactivity for several reasons. First, finger tapping tests are widely used in the research of various neurological conditions (such as amyotrophic lateral sclerosis), for characterization and quantitative assessment of the motor function (Londral et al., 2016). Second, finger movements, as well as arm, head, and foot, are used to assess speed and quality deficits in motor activity of ADHD (Carte, Nigg, & Hinshaw, 1996; Schuerholz, Cutting, Mazzocco, Singer & Denckla, 1997). For instance, finger-sequencing task was recently used by Gaddis et al. (2015), as a measure of motor overflow in children with ADHD. ADHD children display increased motor activity than controls in low stimulation conditions (Antrop, Roeyers, Van Oost, & Buysse, 2000; Kofler, Raiker, Sarver, Wells, & Soto, 2016).

Thus, measuring motor activity after the original response was performed and no novel information was presented, which allow the test to evaluate increased activity precisely in such low stimulation conditions.

By separating commission errors due to impulsive behavior from commission errors due to motor hyperresponsivity, it was possible to identify the multiple sources of response disinhibition. Score in this index was calculated as the average of hyperactive responses along the eight test levels.

**Procedure**

The test was administered by a technician who ensured that the participant understood and followed the instructions. The technician was present throughout the entire session. All participants (including the ADHD group) did not use any medication before or during their participation in the study.

**Data Analysis**

The usefulness of the test for ADHD diagnosis was examined by comparing data along the four MOXO-CPT indices between ADHD and controls. First, Chi-square analysis was used to examine group differences in gender distribution. Then, CPT performance of ADHD children was compared to that of typically developing children, using independent samples T-tests with Tukey’s correction for multiple comparisons for each one of the four MOXO-CPT indices. Also, a total score of the test (consisted of all four indices together) was calculated by averaging the z-scores of the four single indices.

The diagnostic value of the MOXO-CPT was assessed by calculating the areas under the receiver operating characteristic (ROC) curves, which was used to assess the best cutoff points to distinguish between ADHD and non-ADHD children. ROC curves give an indication of the diagnostic ability of a measure to distinguish between the cases and the controls. They yield an area under the curve statistic (ROC-AUC), which ranges from 0 to 1. An ROC-AUC of 0 indicates incorrect classification of all cases, 0.5 indicates that the measure is no better than chance at distinguishing the cases and the controls, and an ROC-AUC of 1 indicates that all participants were correctly classified. Confidence intervals (95%) for an ROC-AUC that do not include 0.5 indicate that a measure can significantly distinguish between the cases and the unaffected controls at the above chance level. Here, MOXO-CPT indices were analyzed for their ability to distinguish between ADHD children and controls. AUC results were considered excellent for AUC values between 0.9 and 1, good for AUC values between 0.8 and 0.9, fair for AUC values between 0.7 and 0.8, poor for AUC values between 0.6 and 0.7, and failed for AUC values between 0.5 and 0.6 (Obuchowski, Lieber, & Wians, 2004). All analyses were conducted with SAS software for Windows version 9.2. A $p$-value of $\leq 0.05$ was considered statistically significant.
Results

Differentiating between ADHD and Typically Developing Children

Results of t-test analyses for independent samples revealed significant differences between ADHD children and their non-affected peers (Table 2). ADHD children received significantly lower scores in the Attention and Timing indices than healthy controls. ADHD children were less attentive to the stimuli, made more omission errors, and performed fewer reactions on accurate timing. Furthermore, ADHD children received significantly higher scores in the Hyperactivity and Impulsivity indices than healthy controls. That is, they performed significantly more commission errors from all types: impulsive commission errors (responding to a non-target stimulus as if it was a target stimulus) and hyperactive commission errors that include all other sorts of non-inhibited responses.

Diagnostic Utility of the MOXO-CPT

To classify children based on the MOXO-CPT performance, the sensitivity and specificity in diagnosis of ADHD were calculated using optimal cutoff as the threshold for indices.

ADHD diagnosis was given according to DSM-IV-TR (APA, 2000).

Table 3 shows the cutoff points, sensitivity, and specificity rates based on the total scores of the MOXO-CPT (taking into account all four indices). Different sensitivity and specificity rates could be used for different purposes of the test (e.g., screening, diagnosis). However, we used the optimal value that represents the maximized classification accuracy with the highest sensitivity and specificity rates. Using the MOXO-CPT, total score revealed that in all age categories, the optimal cutoff points were associated with both high sensitivity and specificity rates. In all age categories, specificity and sensitivity rates were 85% or higher.

The discriminant ability of different MOXO-CPT indices was evaluated by an ROC curve analysis. Table 4 summarizes diagnostic efficiency with the AUC statistic. With the exception of the impulsivity index, all MOXO-CPT indices showed fair to excellent ability to distinguish between the ADHD children and their unaffected peers. AUC values for attention parameter ranged between 0.75 and 0.91, for the timing parameter 0.80 and 0.90, and for hyperactivity 0.73 and 0.82. The impulsivity parameter, however, showed AUC between 0.58 and 0.65, indicating low diagnostic performance when considered alone. The MOXO-CPT total score, which integrates all four CPT indices, consistently showed excellent discriminant ability for all observed ages, with AUC values above 0.91. Figure 4 exemplifies this pattern in children aged 8 years.

Discussion

This study evaluates the usefulness of the MOXO-CPT (Berger & Goldzweig, 2010) in the diagnosis of ADHD in children aged 7–12 years. In line with the previous CPT literature, results showed that the MOXO-CPT consistently distinguished between children with ADHD and their unaffected peers, so that children with ADHD performed worse than controls in all study indices (Corkum & Siegel, 1993; Losier, McGrath, & Klein, 1996; Uno et al., 2006). The largest difference between ADHD and control children was found in the total score of the test that integrates all four indices. As compared to healthy children, children with ADHD showed fewer accurate responses (more omission errors) and fewer reactions on accurate timing. The findings that ADHD children demonstrated more inattentive responses both in quality and in timing provide evidence that RT in CPT may reflect multiple interacting processes (Huang-Pollock et al., 2012) and highlight the need to separately address those two different deficiencies in ADHD.

Table 2. Group differences in MOXO-CPT performance indices

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Difference (ADHD – control)</th>
<th>Attention</th>
<th>Timing</th>
<th>Hyperactivity</th>
<th>Impulsivity</th>
<th>Total score</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>t(124) = −8.75***</td>
<td>t(124) = −10.66***</td>
<td>t(124) = 4.40***</td>
<td>t(124) = 2.51*</td>
<td>t(124) = 11.01***</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>t(140) = −7.79***</td>
<td>t(140) = −11.26***</td>
<td>t(140) = 6.14***</td>
<td>t(140) = 2.29*</td>
<td>t(140) = 11.91***</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>t(136) = −5.47***</td>
<td>t(136) = −8.04***</td>
<td>t(136) = 4.91***</td>
<td>t(136) = 2.90**</td>
<td>t(136) = 10.44***</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>t(147) = −5.12***</td>
<td>t(147) = −8.25***</td>
<td>t(147) = 6.21***</td>
<td>t(147) = 3.76***</td>
<td>t(147) = 11.69***</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>t(136) = −5.17***</td>
<td>t(136) = −7.69***</td>
<td>t(136) = 6.38***</td>
<td>t(136) = 3.82***</td>
<td>t(136) = 12.18***</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>t(112) = −4.50***</td>
<td>t(112) = −6.19***</td>
<td>t(112) = 4.40***</td>
<td>t(112) = 3.29**</td>
<td>t(112) = 8.71***</td>
<td></td>
</tr>
</tbody>
</table>

ADHD, attention-deficit hyperactivity disorder. *p < .05; **p < .01; ***p < .001.
Moreover, children with ADHD showed significantly more commission errors from different types: impulsive commission errors (responding to a non-target stimulus as if it was a target stimulus) and hyperactive commission errors that include all other sorts of non-inhibited responses. Halperin and colleagues (1991) argued that CPT commission errors can stem from multiple sources, which can be distinguished based on the specific nature of the errors. Using the A–X CPT, they noted that commission errors associated with rapid false responses are generally associated with impulsivity, whereas those associated with delayed responses (i.e., X-only errors) are typically associated with inattention. In line with this view, this study suggests that commission errors do not comprise a unitary measure and may reflect differential underlying psychological processes. It supports a multidimensional model of impulsivity and points to the importance of using multiple measures to assess its interrelated components (Lane, Cherek, Rhoades, Pietras & Tcheremissine, 2003; Reynolds, Ortengren, Richards & de Wit, 2006).

Results of the ROC analyses showed that although attention, timing, and hyperactivity indices showed fair to excellent ability to distinguish between ADHD children and controls, the impulsivity index showed poor ability of doing so. Although poor diagnostic capacity of commission errors was previously documented in other CPTs (Fazio, Doyle, & King, 2014), here, it may be attributed to the high cognitive load of the task. In contrast to other CPTs, the present task included environmental distracters that may increase the complexity of the task for both study groups. These higher cognitive demands, which are known to increase commission error rates (Seymour, Mostofsky & Rosch, 2016), may explain why this index is not sensitive and specific enough to ADHD. Further research is needed to elucidate how the index’s clinical utility is affected by different cognitive demands (e.g., with or without distractors).

### Table 3. Specificity and sensitivity rates of the MOXO-CPT total scores

<table>
<thead>
<tr>
<th>Age category (years)</th>
<th>Optimal Cutoff</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>184.60</td>
<td>86</td>
<td>85</td>
</tr>
<tr>
<td>8</td>
<td>167.98</td>
<td>91</td>
<td>88</td>
</tr>
<tr>
<td>9</td>
<td>144.98</td>
<td>90</td>
<td>85</td>
</tr>
<tr>
<td>10</td>
<td>110.88</td>
<td>81</td>
<td>85</td>
</tr>
<tr>
<td>11</td>
<td>107.89</td>
<td>86</td>
<td>89</td>
</tr>
<tr>
<td>12</td>
<td>91.82</td>
<td>85</td>
<td>85</td>
</tr>
</tbody>
</table>

AUC, area under the curve; ROC, receiver operating characteristic curve.

### Table 4. Area under the curve from receiver operating characteristic curve analyses (+95% confidence intervals) for MOXO-CPT indices

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>ROC-AUC</th>
<th>Attention</th>
<th>Timing</th>
<th>Hyperactivity</th>
<th>Impulsivity</th>
<th>Total score</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>0.91</td>
<td>0.86–0.96</td>
<td>0.90</td>
<td>0.85–0.96</td>
<td>0.73</td>
<td>0.64–0.82</td>
</tr>
<tr>
<td>8</td>
<td>0.85</td>
<td>0.76–0.92</td>
<td>0.91</td>
<td>0.86–0.96</td>
<td>0.78</td>
<td>0.71–0.86</td>
</tr>
<tr>
<td>9</td>
<td>0.77</td>
<td>0.68–0.85</td>
<td>0.85</td>
<td>0.77–0.92</td>
<td>0.74</td>
<td>0.65–0.83</td>
</tr>
<tr>
<td>10</td>
<td>0.74</td>
<td>0.66–0.83</td>
<td>0.82</td>
<td>0.75–0.89</td>
<td>0.78</td>
<td>0.71–0.86</td>
</tr>
<tr>
<td>11</td>
<td>0.78</td>
<td>0.70–0.87</td>
<td>0.82</td>
<td>0.74–0.90</td>
<td>0.80</td>
<td>0.72–0.88</td>
</tr>
<tr>
<td>12</td>
<td>0.75</td>
<td>0.65–0.85</td>
<td>0.80</td>
<td>0.71–0.89</td>
<td>0.82</td>
<td>0.73–0.91</td>
</tr>
</tbody>
</table>

AUC, area under the curve; ROC, receiver operating characteristic curve.
Our results showed that the total score of the MOXO-CPT yielded the highest sensitivity and specificity rates, as compared to any single index. MOXO-CPT’s total score was highly accurate in identifying participants with ADHD, based on the DSM-IV-TR (APA, 2000) criteria. For all ages, optimal cutoff values were associated with both high sensitivity and specificity rates of 85% or more. AUC values of the MOXO-CPT total score were excellent, indicating a very high diagnostic performance in all observed ages. Comparing to AUC values of single MOXO-CPT indices, the total score had the greatest ability to discriminate between ADHD children and controls.

These results lend support to the notion that most CPT tasks require multiple cognitive abilities (Coull, Frith, Frackowiak, & Grasby, 1996; Hall et al., 2016; Shalev, Ben-Simon, Mevorach, Cohen & Tsal, 2011; Straube, Bischoff, Nisch, Sauer, & Volz, 2002), and therefore an integration of CPT indices may better reflect the complexity and heterogeneity of ADHD etiology and clinical manifestations.

Several limitations of this study should be considered. The first limitation results from the study’s sampling method. Participation in the study was based on a voluntary agreement of children and their parents. This self-selected sampling strategy tends to be biased towards favoring more cooperative and motivated individuals. Therefore, it is not possible to determine whether this sample also represents other children that were not recruited and whether cooperation is confounded with ADHD variables (Gau et al., 2010; Lee & Ousley, 2006). Moreover, the clinics from which ADHD children were recruited were based in tertiary care hospital. This population has heterogeneous background characteristics including correlates of ADHD. However, the fact that the control group was recruited from a random population supports our findings by showing that the test is able to identify the ADHD children from a random population sample.

Another limitation of the study is the exclusion of ADHD children with severe comorbidities. ADHD is associated with many psychiatric disorders (Gentile, Atiq, & Gillig, 2006), and current CPTs fail to distinguish between patients with ADHD and those with other psychiatric states (Advokat, et al., 2007; Suhr, Sullivan, & Rodriguez, 2011). Choosing a control group of children without comorbidities reflects the study’s focus on the ability of the MOXO-CPT to differentiate ADHD children from their non-affected peers. Nevertheless, it clearly limits our knowledge about the test’s ability to differentiate ADHD from other comorbid conditions. This utility should be further examined. Finally, selection of a threshold for a screening test is best achieved according to the needs of the specific setting in which it is to be used. “Optimal” cutoff values vary depending on the risk–benefit ratio between false-positive and false-negative test results and the base rate of the target disorder in the population at hand. Nevertheless, important information may be lost when defining sensitivity and specificity in relation to a single cutoff value of a continuous variable (Sox, 1986).

Overall, the results of this study indicate several strengths of the MOXO-CPT. The test is able to provide important information in assessing the core variables in ADHD, especially of attention, timing, activity, and impulsivity and reduce reliance on subjective observer reports. Although symptoms of ADHD change considerably during maturation (Berger, Slobodin, Aboud, Melamed, & Cassuto, 2013; Faraone, Biederman, & Mick, 2006), the test’s
ability to discriminate between ADHD children and typically developed controls was consistent in all age group. Specifically, the MOXO-CPT did not show age-related ceiling effect, suggesting that task demands are sufficient even for older children.

On the technical level, participants did not demonstrate any difficulties understanding the instructions or completing the task. Its short administration time may appeal for professionals working with young children who find it hard to cooperate with longer tasks.

There is a need for further research in order to explore the psychometric properties of the MOXO-CPT in other age groups, in samples with comorbid features, and in different subtypes of ADHD. Given the shortcomings of current CPT in differentiating between ADHD and psychiatric controls (McGee, Clark, & Symons, 2000), such study is necessary in further research of the clinical utility of the MOXO-CPT in the diagnostic procedure of ADHD.

Acknowledgements

The authors wish to acknowledge the dedicated help of Merav Aboud and Julia Melamed and thank them for their professional, warm, and kind attitude towards the participating children and their families.

Conflict of Interest

IB and OS served on the scientific advisory board of NeuroTech Solutions Ltd. HC declares no conflicts of interests.

References


