Assessment
2014, Vol. 21(1) 108–118

DOI: 10.1177/1073191112443409
asm.sagepub.com
SAGE

Reliability and Validity of the Online Continuous Performance Test Among Young Adults

Sivan Raz[1], Yair Bar-Haim[2], Avi Sadeh[2], and Orrie Dan[1]

Abstract

Continuous Performance Tests (CPTs) are used in research and clinical contexts to measure sustained attention and response inhibition. Reliability and validity of a new Online Continuous Performance Test (OCPT) was assessed. The OCPT is designed for delivery over the Internet, thereby opening new opportunities for research and clinical application in naturalistic settings. In Study 1, participants completed the OCPT twice over a 1-week period. One test was taken at home and one in the laboratory. Construct validity was assessed against a gold standard CPT measure. Results indicate acceptable reliability between the home- and laboratory-administered tests. Modest to high correlations were observed between the OCPT scales and the corresponding scales of the gold standard CPT. Study 2 examined whether the OCPT may discriminate participants with attention deficit hyperactivity disorder from healthy controls. Results revealed significantly higher rates of omission and commission errors and greater response time variability in participants with attention deficit hyperactivity disorder relative to healthy controls. These results support the reliability and validity of the OCPT and suggest that it may serve as an effective tool for the assessment of attention function in naturalistic settings.

Keywords

continuous performance tests, CPT, online CPT, ADHD, reliability, validity

Continuous Performance Tests/Tasks (CPTs) require subjects to maintain vigilance and react to the presence (or absence) of a specific stimulus within a set of continuously presented distracters (McGee, Clark, & Symons, 2000). CPTs have a long history of use in measuring processes related to vigilance, sustained attention, response inhibition, and other aspects of attention and cognitive function (Conners, Epstein, Angold, & Klaric, 2003; Riccio, Reynolds, & Lowe, 2001). CPTs have been extensively used in the context of learning and memory research (e.g., Pastor & Reuben, 2008; Riccio et al., 2001), sleep restriction and deprivation studies (e.g., Goel, Rao, Durmer, & Dinges, 2009; Sadeh, Dan, & Bar-Haim, 2011), and on a wide variety of psychiatric conditions such as schizophrenia, bipolar disorder, dementia, and conduct disorder (e.g., Advokat, Martino, Hill, & Gouvier, 2007; Bora, Yucel, & Pantelis, 2009; Buchanan, Strauss, Breier, Kirkpetrick, & Carpenter, 1997; O'Brien et al., 1992), and in particular in attention deficit hyperactivity disorder (ADHD; e.g., Barkley, 1991, 1998; Corkum & Siegel, 1993; Epstein et al., 2003; Losier, McGrath, & Klein, 1996; Nichols & Waschbusch, 2004; Riccio & Reynolds, 2001).

Several norms-based CPTs are available as hardware–software packages (e.g., Conners's Continuous Performance Test [CPT-II], Integrated Visual and Auditory CPT [IVA], Test of Variables of Attention [TOVA]). The extant CPTs are designed for administration under strictly controlled conditions and sometimes require specifically designated hardware. Although such requirements provide important, and some may argue necessary, experimental control over task procedures, they also restrict the generalizability of the findings and do not allow for mass testing of populations in naturalistic settings. Thus, a reliable and valid CPT test that can be taken online may hold noticeable practical advantages for field research and clinical follow-up purposes. First, an online CPT may be more cost effective than standard CPTs performed in the laboratory or the clinic. Second, an online test may enhance dissemination of such specialized testing because tests could be taken in a variety of settings (e.g., home, work place, school), at different times during the day, and by multiple persons at the same time. Third, and akin with the previous point, an online CPT may

[1]The Center for psychobiological Research, Department of Psychology, The Max Stern Yezreel Valley College
[2]Department of Psychology, Tel Aviv University

Corresponding Author:
Sivan Raz, Department of Psychology, The Center for Psychobiological Research, The Max Stern Academic College, Emek Yezreel 19300, Israel.
Email: sivanr@yvc.ac.il

prove particularly useful for the assessment of children, elderly, disabled persons, and others who experience difficulties in traveling or adjusting to the demands associated with laboratory settings. Finally, taking the CPT test online, from home or from the office, can contribute to a much higher compliance rate. All these advantages notwithstanding, one must still consider the relative lack of control over the testing environment that might interfere with the reliability and validity of test results.

Therefore, the purpose of the present study was to assess the reliability and validity of a new Online Continuous Performance Test (OCPT), for use as an ambulatory assessment tool of sustained attention, response inhibition, and response time consistency. To our knowledge, the OCPT is the first continuous performance task designed and programmed for delivery over the Internet. The OCPT is conceptually and operationally similar to other CPTs described in the literature and therefore relies on both face validity and on prior research with similar established tasks. We expected the reliability and validity of the OCPT to be comparable to those of similar more traditional CPTs such as Conner's CPT-II and the TOVA.

The present studies evaluate split-half reliability, test–retest reliability, and construct validity of the OCPT.

Study 1

In Study 1, split-half reliability (internal consistency) of the OCPTs taken at home and under standardized conditions in the laboratory was examined. In addition, 1-week test–retest reliability was evaluated with one test taken at home and the other taken at the laboratory. The OCPT's construct validity was tested against a gold standard CPT (Conners's CPT-II; Conners et al., 2003; Egeland & Kovalik-Gran, 2010; Epstein et al., 2003; Hill, Pella, Singh, Jones, & Gouvier, 2009; McGee et al., 2000; Rabin, Barr, & Burton, 2005). We expected (a) adequate split-half coefficients for all the derived indices of the OCPT, (b) significant correlations between the home-taken OCPT and the laboratory-taken OCPT, and (c) significant correlations between both the home-and laboratory-taken OCPTs and the corresponding indices of Conners's CPT-II taken in the lab. The conventionally accepted levels of reliability in psychometric studies are 0.7 at the early stages of construct validation research, 0.8 for tests in established areas, and 0.9 for tests that will be used in making decisions about individuals (Nunnally & Bernstein, 1994). However, previous studies examining the reliability of continuous performance tests usually report somewhat lower reliability levels ranging between 0.4 and 0.7, typically considered as moderate (Barkley, 1991; Conners, 2004; Forbs, 1998; Leark, Wallace, & Fitzgerald, 2004; Llorente et al., 2001; Llorente et al., 2008; Nichols & Waschbusch, 2004). We therefore expected the OCPT reliability results to be in line with these other CPTS.

Method

Participants. Thirty-nine undergraduate students (mean age = 24.47 ± 4.62 years, range = 19-33 years; 29 females) participated in the study. Participants received course credit according to their academic requirements. All participants were reportedly in good health and taking no medications or drugs. Written informed consent was obtained from all participants. The study was approved by the institutional review board.

Measures

Online Continuous Performance Test. The OCPT (eAgnosis Inc., Newark, DE) is a standard CPT designed and programmed for delivery over the Internet. The task is adapted for use with most Internet browsers (e.g., Internet Explorer, Chrome) and can be operated from practically any computer running a Windows 98 operating system or higher that is connected to the Internet (for a demo, see www.check-adhd.com/onlineCPTresearch.php). Milliseconds accuracy in response time (RT) recording is achieved by using the end user's own CPU and transmitting the collected data back to a remote database at various points during task performance. Thus, the integrity of data collection does not dependent on Internet speed or bandwidth. Furthermore, because the programming code for RT monitoring and recording is compact, the end user's computer characteristics and processor speed do not affect recording accuracy.

The task uses two geometric stimuli: equilateral triangles (~5 cm sides) and circles (~5 cm diameter), both presented in a light blue color delineated by a black one-pixel stroke contour. The geometric stimuli are presented in the middle of the screen against a gray background within a 4-pixle stroke 7.5 × 11 cm black rectangle that is presented constantly throughout the task.

Each trial consists of a presentation of one geometric shape for 100 ms followed by a 1,900 ms intertrial interval. Participants are instructed to respond to the triangle shape as fast as possible without compromising accuracy by pressing the spacebar on the computer's keyboard and to withhold response to the circle shape. The test consists of two conditions: low target frequency and high target frequency (see Greenberg, Kindschi, Dupuy, & Corman, 1996; Greenberg & Waldman, 1993, for a similar task presentation). The first half of the test (the low target frequency half), consists of 224 trials (56 targets, 168 nontargets) with a target to nontarget ratio of 1:3. This half of the test is boring and fatiguing. In the second half of the test (the high target frequency half), the target to nontarget ratio is reversed and is set to 3:1 (168 targets, 56 nontargets). In this half of the test, the participant expects to respond most of the time but occasionally must inhibit the tendency to respond. The manipulation of target frequency in CPTs is designed to highlight the sustained attention component in the low target frequency condition and the

response inhibition component in the high target frequency condition. Parameters of the low target frequency condition of the OCPT were modeled after those of previously shown effective indices of sustained attention (e.g., the TOVA; Greenberg & Waldman, 1993).The high frequency condition was selected to places enhanced demands on response inhibition and impulse control by inducing a strong response set. Thus, the varying target to nontarget ratio allows users of the OCPT to examine the effects of differing response demands or response sets on inattention and impulsivity (Greenberg & Waldman, 1993). Each half of the OCPT is further subdivided into four blocks, each of which contains an identical number of trials and target stimuli, thus preserving a balanced target/nontarget presentation ratio across blocks. Both the high and low target conditions of the OCPT involve a go–no-go element.

Throughout the task the geometric shapes are presented in a fixed, pseudo random sequence. To minimize practice effects, each half of the test is preceded by a 2-minute practice phase reflecting the target to nontarget ratios of the actual test to follow. Subjects are not informed about the practice nature of these 2-minute practice sessions and consider these as part of the test. Total net test time of the OCPT (including the two practice sessions) is 19 minutes. Three breaks are allowed (following the first and second practice sessions and following the low target frequency session). Participants are instructed not to exceed 2 minutes of break time. Responses with RTs that are faster than 150 ms are considered anticipatory and are removed from analyses. When a participant presses the spacebar more than once per stimulus presentation, only the data from the first response is included in data analysis.

The software automatically records the participant's omission and commission errors as well as reaction times. Four primary measures are extracted for analyses: errors of omission (defined as the number of targets to which a participant did not respond), errors of commission (defined as the number of times a participant incorrectly responded to a nontarget), response times, and response time consistency (standard deviations of response times). These measures are extracted per condition (low and high target frequencies) and can also be extracted by test blocks (quartiles) per condition. The software alerts if an unusual pattern of results is detected (i.e., considerable amount of missing data during significant portions of the test). All data are stored offline and are available as a Microsoft Excel spreadsheet.

Conners's Continuous Performance Test. Conners's CPT-II (Conners, 2004) is a computerized visual vigilance test of attention. The test takes 14 minutes to complete and consists of 360 presentations of letters (1 inch in size) that appear on the screen one at a time for 250 ms. Letters are presented in 18 consecutive blocks of 20 trials each. Participants are required to press the spacebar when any letter other than the letter "X" appears on the screen. The letter "X" appears on 10% of all trials. As with the OCPT, four primary performance indices are derived: an index of response inhibition (the percentage of commission errors—i.e., false positives) defined as the number of times the participant incorrectly responded to an X; an index of vigilance (percentage of omission errors—targets to which the participant did not respond); hit rate reaction time; and an index of response time consistency defined as the standard error of hit rate reaction time. Test–retest reliabilities of the CPT-II over a 3-month interval range between 0.55 and 0.84, and split-half coefficients range between 0.73 and 0.95 (Conners, 2004). The CPT-II is considered a gold standard measure for sustained attention and response inhibition and was therefore chosen to serve as an external measure against which the construct validity of the OCPT is assessed. The CPT-II task most closely resembles the OCPT's high target frequency condition.

Procedure. Participants were invited to the laboratory where they were instructed how to access the OCPT from home and how to complete the test. They were instructed to take the home test on the same day and time as their prescheduled laboratory-based retest one week later. Participants were also requested to take the test in a quiet room and to avoid potential distracters and interruptions. One week after completion of the home OCPT, participants came to the laboratory and completed Conners's CPT-II and the OCPT in a quite testing room. A 20-minute break was taken between the two lab-taken tests, and order of testing was counterbalanced across participants.[1]

Results

Means and standard error means (*SEM*) of omission errors, commission errors, RTs, and response time consistency of the home-taken OCPT and the laboratory-taken OCPT and CPT-II are presented in Table 1. Intercorrelations between low and high target frequency sections for the home and laboratory OCPTs are presented in Table 2.

OCPT split-half reliability. To examine internal reliability of the OCPT, split-half analyses were conducted on both home and laboratory taken tests. Specifically, split-half correlations between mean omission errors, mean commission errors, mean reaction times, and mean response consistencies on Blocks 1and 4 and Blocks 2 and 3 of the low and high target frequency portions of the OCPT were conducted. The Spearman–Brown formula was used to adjust the split-half correlation (Halperin, Sharma, Greenblatt, & Schwartz, 1991). Table 3 shows correlations matrices representing internal consistency coefficients for omissions, commissions, RT, and response consistency. The split-half correlation coefficients for all measures in the high target condition showed high levels of reliability both at home (range = 0.83-0.98) and at the laboratory (range = 0.84-0.97). Split-half coefficients were in the moderate-to-high range for all

Table 1. Means and Standard Errors (SEMs) for Omission Errors, Commission Errors, Reaction Times (ms), and RTSD for the Home- and Laboratory-Taken OCPTs and the CPT-II.

	OCPT Home				OCPT Laboratory					
	Low Target		High Target		Low Target		High Target		CPT-II	
	Mean Value (*SEM*)	Mean Percentile (*SEM*)	Mean Value (*SEM*)	Mean Percentile (*SEM*)	Mean Value (*SEM*)	Mean Percentile (*SEM*)	Mean Value (*SEM*)	Mean Percentile (*SEM*)	Mean Value (*SEM*)	Mean Percentile (*SEM*)
Omissions	1.15 (0.28)	2.07 (0.51)	2.89 (0.61)	1.74 (0.37)	0.82 (0.21)	1.47 (0.38)	2.79 (0.61)	1.72 (0.37)	1.51 (0.42)	0.46 (0.13)
Commissions	2.05 (0.75)	1.23 (0.45)	9.30 (1.56)	16.49 (2.87)	1.17 (0.29)	0.71 (0.17)	7.28 (1.48)	13.18 (2.71)	11.17 (1.27)	31.12 (3.53)
RT	464.51 (11.70)		407.99 (11.84)		445.18 (11.91)		373.19 (11.69)		384.38 (12.74)	
Response consistency[a]	84.06 (3.81)		107.08 (9.09)		83.14 (5.19)		92.21 (6.68)		5.38 (0.32)	

Note. OCPT = Online Continuous Performance Test; *SEM* = standard error of mean; RTSD = response time consistency; RT = response time.
a. The OCPT response consistency is calculated with standard deviation values whereas the CPT-II response consistency calculated with standard error values and therefore the differences in magnitude.

Table 2. Intercorrelations Between Low and High Target Frequency Conditions of the Home and Laboratory OCPTs for Omission Errors, Commission Errors, Reaction Time (ms), and RTSD.

		Home, High Target				Laboratory, High Target			
		Omissions	Commissions	RT	Response Consistency	Omissions	Commissions	RT	Response Consistency
Low target	Omissions	**0.33***	0.18	−0.13	0.04	**0.44****	0.24	0.13	0.50***
	Commissions	0.65***	**0.56*****	−0.29	0.68***	0.46**	**0.59*****	−0.20	0.28
	RT	−0.21	−0.47***	**0.81*****	0.02	−0.07	−0.40**	**0.85*****	0.40**
	Response consistency	0.17	−0.06	0.08	**0.13**	0.28	−0.08	0.34*	**0.48*****

Note. OCPT = Online Continuous Performance Test; RTSD = response time consistency; RT = response time.
In bold are correlations between congruent measures in the high and low target conditions.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table 3. Split-Half Consistency (Q1&4 vs. Q2&3 in the LT and HT Frequency Sections) for Errors of Omission, Errors of Commission, Response Time (ms), and RTSD at Home and Laboratory.

	Home				Laboratory			
	LTQ1&4	LTQ2&3	HTQ1&4	HTQ2&3	LTQ1&4	LTQ2&3	HTQ1&4	HTQ2&3
Omissions								
LTQ1&4	1				1			
LTQ2&3	**0.66**	1			**0.63**	1		
HTQ1&4	0.45	0.16	1		0.18	0.71	1	
HTQ2&3	0.60	0.45	**0.83**	1	0.29	0.75	**0.91**	1
Commissions								
LTQ1&4	1				1			
LTQ2&3	**0.93**	1			**0.74**	1		
HTQ1&4	0.66	0.76	1		0.74	0.73	1	
HTQ2&3	0.62	0.69	**0.95**	1	0.60	0.62	**0.93**	1
RT								
LTQ1&4	1				1			
LTQ2&3	**0.97**	1			**0.97**	1		
HTQ1&4	0.90	0.89	1		0.91	0.90	1	
HTQ2&3	0.87	0.86	**0.98**	1	0.92	0.91	**0.97**	1
Response consistency								
LTQ1&4	1				1			
LTQ2&3	**0.57**	1			**0.66**	1		
HTQ1&4	0.32	0.31	1		0.64	0.62	1	
HTQ2&3	0.27	0.25	**0.88**	1	0.60	0.59	**0.84**	1

Note. LT = low target; HT = high target; RT = response time; HTQ = high target quarters; LTQ = low target quarters; RTSD = response time consistency.
In bold are correlations between congruent measures in the 1st+4th quartiles and the 2nd+3rd quartiles.

Table 4. Intraclass Correlation Coefficients Between Home-Taken and Laboratory-Administered OCPTs for Omission Errors, Commission Errors, Reaction Time (ms), and RTSD.

	Laboratory							
	Low Target				High Target			
	Omissions	Commissions	RT	Response consistency	Omissions	Commissions	RT	Response consistency
Home								
Low target								
Omissions	**0.20**	0.11	−0.07	0.02	**0.24**	−0.01	0	0.01
Commissions	0.09	**0.24**	−0.04	0.01	0.46**	**0.51*** **	−0.04	0.09
RT	0	0	**0.52*** **	0.14	−0.02	−0.12	**0.48*** **	0.04
Response consistency	0.02	0.02	0.19	**0.36****	0.09	−0.09	0.17	**0.23**
High target								
Omissions	**0.14**	0.21	−0.02	0.03	**0.39****	0.30	−0.03	0.05
Commissions	0.06	**0.22**	−0.13	−0.02	0.29*	**0.83*** **	−0.14	0.07
RT	0	−0.01	**0.50*** **	0.14	−0.02	−0.12	**0.58*** **	0.13
Response consistency	0.02	0.02	−0.02	**0.16**	0.05	0.14	−0.08	**0.51*** **

Note. OCPT = Online Continuous Performance Test; RT = response time.
*$p < .05$. **$p < .01$. ***$p < .001$.

measures in the low target condition both at home (range = 0.57-0.97) and at the laboratory (range = 0.86-0.97). In the low target condition, the indices of commission errors and RT tended to exhibit greater consistency relative to the indices of omission errors and response time consistency.

Home versus laboratory performance on the OCPT (test–retest reliability). To assess differences in performance between tests taken at home and tests taken in the laboratory, paired sample *t* tests were computed. No significant differences were observed between home and laboratory performance on any of the task indices in the low target frequency section of the OCPT (all *p*s > .10). In the high target frequency section of the OCPT, no significant differences were observed between home and laboratory performance in omission errors ($p > .90$) and response time consistency ($p > .07$). However, more commission errors were detected at home (9.30 ± 1.56) compared with the laboratory (7.28 ±1.48), $t(38) = 2.27$, $p < .03$, and mean RT was slower at home (408 ± 12 ms) than at the laboratory (373 ± 12 ms), $t(38) = 3.24$, $p < .002$.

To further examine the correspondence between OCPT performances at home and at the laboratory, as well as to assess 1 week test–retest reliability of the OCPT, intraclass correlation coefficients were computed among the different OCPT measures, separately for the high and low target frequency portions of the OCPT (Table 4). These analyses indicate modest to strong correspondence between the results of the tests taken at home and in the lab, as well as acceptable test–retest reliability.

Validation of the OCPT against the CPT-II (construct validity). Correlations between the OCPT measures at home and in the laboratory and the CPT-II scores were examined using Pearson correlations, separately for the high and low target frequency portions of the OCPT (Table 5). Except for the correlations between the home-taken OCPT omission errors index in the low target section and the response consistency index in the high target section and the corresponding CPT-II indices, all other correlations between corresponding measures of the OCPT and the CPT-II were statistically significant showing modest to strong associations (range = 0.40-0.79).

Differences in the magnitude of correlations between the CPT-II and the corresponding scales of the OCPT taken at the laboratory relative to home environment were tested using Fisher's *r*-to-*z* transformation. Analysis indicated that those differences were not significant (all *p*s > .10) except for the RT measure in both the low ($p = .006$) and high ($p = .05$) target portions of the OCPT.

Discussion

Reliability. One advantage offered by OCPT is the capacity to collect data in naturalistic environments. It was therefore important to determine whether the conditions under which the test is taken affect performance patterns. The results indicate that despite the relatively small sample size, the wide variability in the conditions under which the OCPT was taken during the home session, and the lack of direct experimental supervision over the home OCPT sessions, the home and laboratory results show acceptable correspondence in participants' performance. Out of eight primary expected correlations, only two failed to reach statistical significance (omission and commission errors at the low target frequency section of the OCPT). The lack of

Table 5. Correlations Between Home- and Laboratory-Taken OCPT and CPT-II: Omission Errors, Commission Errors, Reaction Time (ms), and RTSD.

		CPT-II			
		Omissions	Commissions	RT	Response Consistency
Home OCPT					
LT	Omissions	**−0.03**	0.27	−0.07	−0.03
	Commissions	0.62***	**0.43****	−0.34*	−0.12
	RT	−0.14	−0.39*	**0.40****	0.30
	Response consistency	0.15	−0.06	0.38*	**0.50*****
HT	Omissions	**0.49****	0.53***	−0.15	0.05
	Commissions	0.54***	**0.79*****	−0.61***	−0.24
	RT	−0.25	−0.43**	**0.44****	0.26
	Response consistency	0.67**	0.37*	−0.10	**0.23**
Laboratory OCPT					
LT	Omissions	**0.21**	0.25	−0.05	0.27
	Commissions	0.47**	**0.51*****	−0.30	0.03
	RT	0.13	−0.53***	**0.79*****	0.70***
	Response consistency	0.24	0.06	0.30	**0.59*****
HT	Omissions	**0.45****	0.54***	−0.19	0.28
	Commissions	0.61***	**0.73*****	−0.59***	−0.14
	RT	−0.04	−0.50***	**0.73*****	0.58***
	Response consistency	0.60***	0.17	0.15	**0.55*****

Note. CPT = Continuous Performance Test; LT = low target frequency; HT = high target frequency; OCPT = Online Continuous Performance Test; RT = response time.
In bold are correlations between congruent measures of the OCPT taken at home and in the lab and the CPT-II.
*$p < .05$. **$p < .01$. ***$p < .0001$.

significant correlations on these specific indexes of the OCPT may be attributed to the overall low occurrence of omission and commissions errors under this condition (low target) in general.

The results also indicate that test–retest reliability is somewhat stronger for the high target frequency section of the OCPT relative to the low target frequency section of the test. Based on these findings we preliminarily conclude that home and laboratory OCPT results show acceptable, moderate-to-high correspondence (especially with respect to the high target frequency condition), rendering the OCPT a reasonably reliable tool for ambulatory assessment of attention functions in adults.

Comparisons between home and laboratory mean performance scores revealed no differences between the tests in the low target frequency section of the OCPT, where the task is boring and fatiguing. In the high target frequency section of the task, where response inhibition is challenged to a greater extent, some differences between home and laboratory performance were observed. Specifically, more commission type errors and longer reaction times were recorded for the home-taken OCPT relative to the laboratory-administered test. Practice effects might account for these differences in the present study because the laboratory tests were administered after the home tests. However, the results from a separate sample in which the order of task presentations was reversed indicate no order effects on OCPT performance, thereby alleviating this concern to some extent. More research applying a fully counterbalanced task presentation is needed to fully resolve this issue.

Alternately, and perhaps more plausibly, the different settings in which the tests were taken may have affected performance. For instance, the laboratory environment and the presence of an experimenter could have affected motivational and emotional aspects of task performance that contributed to the somewhat better results. To establish a more precise index of the OCPT's temporal reliability, it could be administered twice under the same environmental settings. In such a case it is likely that retest correlations would be higher than those reported in the current study. Analysis of the intercorrelations between the low and high target frequency conditions of the OCPT revealed consistent pattern across home and laboratory administrations. One index, response consistency (RTSD), appeared to be higher for the laboratory relative to the home administration. However, Fisher *r*-to-*z* transformation indicated that this difference was not statistically significant.

Reliability of the OCPT was also demonstrated by the satisfactory split-half coefficients on all measures of the test (Conners, 2004; Llorente et al., 2001; Llorente et al., 2008). Results suggest that when used with young adults the internal consistency of the OCPT is high for the high target

condition of the OCPT and moderate to high for the low target condition of the OCPT. Although the increase in correlation magnitude of the high relative to the low target condition might be attributed to practice effects or the larger number of targets in the high target condition, this was not the case for all correlations. Instead, this finding might have emerged as an artifact of a few extreme outliers. It is also important to note that the split-half coefficients were almost identical for the home and laboratory OCPT administrations.

Validity. Conners's CPT-II was selected as a validation tool for the OCPT because it is considered a gold standard CPT measure, producing valid and reliable assessments of attention function (e.g., Conners et al., 2003; Egeland & Kovalik-Gran, 2010; Epstein et al., 2003; Hill, Pella, Singh, Jones, & Gouvier, 2009; McGee et al., 2000; Rabin, Barr, & Burton, 2005). The present results show significant moderate-to-strong correlations between the primary CPT-II indicators (errors of omission, errors of commission, RT, and response time consistency) and the corresponding OCPT scales both when measured at home and in the laboratory, with the exception of omission errors in the low target section of the OCPT and response consistency in the high target section of the OCPT in the laboratory administration.

There are differences in task demands and stimuli between the OCPT and the CPT-II. The CPT-II is primarily based on response inhibition in the context of long streams of distracters, which is highly compatible with the high target condition of the OCPT. Indeed, the strongest correlations between the two tasks were found in this category. The low correlation in the measure of low target omission errors may be attributed to a differing proportion of target stimuli between the CPT-II and OCPT. It is also possible that some correlation between the tests is (partially) attributable to other underlying cognitive, perceptual, and psychomotor dimensions not controlled for in the current study such as IQ, speed of mental processing, and motor reaction time. Future studies may wish to decompose the contribution of such functions to OCPT performance.

The correlations between OCPT and CPT-II corresponding scales appear to be somewhat stronger for the OCPT administered in the lab relative to the OCPT taken at home. However, Fisher's *r*-to-*z* transformation indicated that those differences in the magnitude of correlations were not significant except for the RT measure in both the low and high target portions of the OCPT. The higher correlations in RT between the CPT-II and the OCPT taken at the laboratory may be attributed to a greater similarity in testing conditions in the lab versus home contexts (same room, same computer, same experimenter vs. different rooms, different computers, and no experimenter), as well as to the shorter time interval between test administrations in the lab versus home (20 min vs. 1 week).

Based on the findings we conclude that both home-taken and laboratory-administered OCPTs are valid instruments for measuring pivotal attention functions such as sustained attention and response inhibition.

Study 2

CPTs are extensively used to investigate indices of attention among individuals with ADHD and are often included in ADHD-focused neuropsychological batteries. Studies have shown greater response time variability and higher frequencies of omission and commission errors in patients with ADHD compared with healthy controls (e.g., Conners, 2004; Epstein et al., 2003; Klein, Wendling, Huettner, Ruder, & Peper, 2006). To further assess the construct validity of the home-taken OCPT, and to examine whether the results of Study 1 may generalize to target populations of interest, in Study 2 we administered the OCPT to participants with and without ADHD diagnosis. It was expected that (a) participants with ADHD display more omission and commission errors compared with participants without ADHD and (b) participants with ADHD will show higher response time variability relative to healthy participants without ADHD. Split-half reliability of the OCPT was again assessed separately for the ADHD and control groups.

Method

Participants. Fifty-five participants (48 females), aged 21 to 29 years (mean age = 24.34 ± 2.12 years) were classified into two groups out of a pool of 222 undergraduate students based on their scores on the ADHD Rating Scale–IV (DuPaul, Power, Anastopoulos & Reid, 1998) and a clinical interview. Participants who reported at least six symptoms of either inattention or hyperactivity-impulsivity on the self-report ADHD rating scale were invited to a clinical interview performed by an experienced senior clinical psychologist (OD) to determine ADHD diagnosis ($n = 28$). The control group ($n = 27$) consisted of participants who reported less than three symptoms of either inattention or hyperactivity-impulsivity on the ADHD rating scale. Absence of ADHD in the control group was further verified in a clinical interview. The ADHD and control groups did not differ on mean age, gender, or ethnicity distribution.

Participants received course credit according to their academic requirements. Written consent was obtained from all participants, and the study was approved by the institutional review board.

Measures

ADHD rating scale. The ADHD Rating Scale-IV (DuPaul et al., 1998) is a valid and reliable self-report scale (Kooij et al., 2008) consisting of 18 items describing symptoms of inattentiveness, hyperactivity, and impulsivity on which *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (*DSM-IV*; American Psychiatric Association, 1994), diagnosis of ADHD is based. Chronbach's αs for the present

Table 6. Split-Half Consistency (Q1&4 vs. Q2&3 in the LT and HT Frequency Sections) for Errors of Omission, Errors of Commission, Response Time (ms), and RTSD Separately for the ADHD and Control Groups.

		ADHD				Control			
		LTQ1&4	LTQ2&3	HTQ1&4	HTQ2&3	LTQ1&4	LTQ2&3	HTQ1&4	HTQ2&3
Omissions									
	LTQ1&4	1				1			
	LTQ2&3	**.92**	1			**.52**	1		
	HTQ1&4	.89	.97	1		.66	.59	1	
	HTQ2&3	.90	.98	**.96**	1	.62	.70	**.83**	1
Commissions									
	LTQ1&4	1				1			
	LTQ2&3	**.79**	1			**.66**	1		
	HTQ1&4	.75	.69	1		−.18*	−.11[a]	1	
	HTQ2&3	.73	.46	**.84**	1	−.44*	−.16[a]	**.81**	1
RT									
	LTQ1&4	1				1			
	LTQ2&3	**.96**	1			**.95**	1		
	HTQ1&4	.76	.77	1		.95	.89	1	
	HTQ2&3	77	.74	**.96**	1	.93	.83	**.96**	1
Response consistency									
	LTQ1&4	1				1			
	LTQ2&3	**.85**	1			**.60**	1		
	HTQ1&4	.74	.86	1		.57	.79	1	
	HTQ2&3	.83	.90	**.94**	1	.82	.54	**.73**	1

Note. ADHD = attention deficit hyperactivity disorder; LT = low target; HT = high target; RTSD = response time consistency; Q1&4 = Quarters 1 and 4; Q2&3 = Quarters 2 and 3.

a. The correlation between forms (halves) of the test is negative. This violates reliability model assumptions. Statistics, which are functions of this value, may have estimates outside theoretically possible ranges.

In bold are correlations between congruent measures in the 1st+4th quartiles and the 2nd+3rd quartiles.

sample were .80 for the ADD scale and .74 for hyperactivity and impulsivity scale.

Clinical interview. The clinical interview was conducted by an experienced senior clinical psychologist. Participants were asked about each of the 18 specified symptoms of ADHD according to the *DSM-IV*. For example, "In the last year, that is, since (name of current month last year) did you have trouble with (name symptom 1-18 from the *DSM-IV*?"); if the answer was "no," the interviewer continued to the next symptom. If the answer was "yes," the interviewer asked the following questions: (a) Did you have this trouble at home? (b) Did you have this trouble at work? (c) Did you have this trouble while studying? (d) How old were you when you started to have (name of the specific symptom from the *DSM-IV*)? (e) Does the symptom cause a significant impairment in social, academic, or occupational functioning? A diagnosis of ADHD was assigned according to the criteria specified in the *DSM-IV*.

Procedure. A total of 222 undergraduate students completed the ADHD Rating Scale-IV. Sixty-nine students met the criteria for either the ADHD group or the control group. Those 69 students were then invited to an interview. Based on the interview, 28 students met criteria for ADHD diagnosis and 27 met criteria for the control group. Fourteen students did not meet the criteria for neither the ADHD or control groups and therefore were excluded from the study. One week later, participants were invited again to the lab and were instructed on how to access the OCPT from home and how to complete the test (see Study 1). In addition, they were instructed to abstain from using methylphenidate on the day of testing.

Results

Split-half reliability. Split-half reliability was calculated as described in Study 1. Table 6 shows correlations matrices representing internal consistency coefficients for OCPT indices. The split-half correlation coefficients for all OCPT measures showed high levels of reliability within the ADHD group (range = 0.79-0.96) and moderate-to-high levels of reliability within the control group (range = 0.52-0.96).

OCPT performance among participants with ADHD relative to healthy controls. Differences in OCPT omissions, commissions, RTs, and response consistency were analyzed separately for the low target and high target conditions of the OCPT using *t* tests.

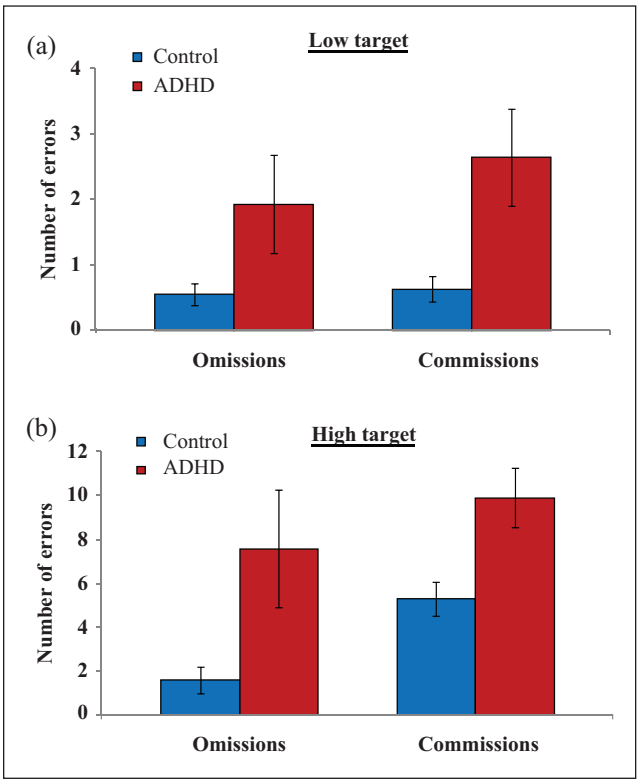

Figure 1. Number of omission and commission errors made by participants with and without ADHD in the low (A) and high (B) target frequency sections of the OCPT.
Note. ADHD = attention deficit hyperactivity disorder; OCPT = Online Continuous Performance Test; *SEM* = standard error means. Bars represent mean, and error bars represent *SEM*.

Low target condition. Participants with ADHD had more commission errors compared with participants without ADHD, $t(53) = 2.57$, $p < .01$; Cohen's $d = 0.70$ (Figure 1A). Participants with ADHD also showed less response consistency (111 ± 16 ms) compared with healthy controls (76 ± 5 ms), $t(53) = -2.10$, $p < .04$; Cohen's $d = 0.57$. A nonsignificant trend was also observed by which participants with ADHD had more omission errors, $t(53) = 1.75$, $p = .08$; Cohen's $d = .48$. No differences between the groups were found in simple RTs.

High target condition. Participants with ADHD had more omission, $t(53) = 2.13$, $p < .04$; Cohen's $d = 0.58$, and commission errors, $t(53) = 2.94$, $p < .005$; Cohen's $d = 0.79$, compared with participants without ADHD (Figure 1B). Participants with ADHD also showed less response consistency (115 ± 13 ms) compared with healthy controls (84 ± 6 ms), $t(53) = 2.14$, $p < .04$; Cohen's $d = 0.58$. No between-group differences were found for RT.

Discussion

Reliability. Split-half analysis revealed satisfactory coefficients on all measures of the OCPT (Conners, 2004; Llorente et al., 2001; Llorente et al., 2008). Results suggest that when used with young adults, the internal consistency of the OCPT is high within participants with ADHD and moderate to high within the control group. The pattern of split-half correlations was similar to that reported in Study 1. This extends the findings of Study 1 by showing that the OCPT taken at home has good internal reliability in a highly relevant clinical population.

Validity. To examine OCPT's contrast validity and its ability to discriminate between participants with different levels of diagnosed attention-related difficulties we administered the OCPT to participants with and without ADHD. Results indicated that OCPT performance was significantly better in terms of number of omission errors, commission errors, and response consistency in the healthy control group relative to the ADHD group. This is in line with other CPT studies (for review, see Klein et al., 2006) reporting significantly higher amount of omission and commission errors and greater response variability in patients with ADHD compared with controls (Epstein, Conners, Sitarenios, & Erhardt, 1998; Epstein, Johnson, Varia, & Conners, 2001; Walker, Shores, Troller, Lee, & Sachdev, 2000). Increased response variability was frequently reported along with normal mean RT, and in almost all studies response variability yielded the strongest between-groups effect sizes. Klein et al. (2006) concluded that increased response variability is thus not secondary to overall slower responding. In the present study, the measure of commission errors yielded the strongest between-groups effect size. Based on the current results, it might be suggested that OCPT measures of response consistency and commission errors contribute best to group discrimination.

The results of this study should be viewed in light of a couple of limitations. First, data in the current study were collected at participants' homes and thus the possibility that patients may be more sensitive to distractions in uncontrolled setting remains unresolved. Future studies may explore differences between ADHD and control participants in OCPT performance at the laboratory as well. Second, although the results of Study 2 support the validity of the OCPT in discriminating participants with and without ADHD, gaining clinical relevance would require additional testing and validation. For example, most of the test–retest correlation coefficients in Study 1 (conducted with a nonclinical sample) fell below the level recommended for clinical decision making (Anastasi & Urbina, 1998). Therefore, test–retest reliability should be further evaluated with larger clinical populations as well as extended to child populations that constitute a major clinical target population in ADHD.

In conclusion, the results of Study 2 indicate good construct validity of the OCPT in differentiating between ADHD and non-ADHD participants by identifying common aspects of attentional/impulsive deficits.

General Discussion

The aim of the present study was to test the reliability and validity of the OCPT. Different measures of reliability and validity were used, including split-half reliability, test–retest reliability, and construct validity. Taken together, these assessments revealed reliability and validity results that may be somewhat lower than those proposed by Nunnally and Bernstein (1994) but are in line with previously published evaluations of CPT tasks (Barkley, 1991; Conners, 2004; Forbes, 1998; Leark, Wallace, & Fitzgerald, 2004; Llorente et al., 2001; Llorente et al., 2008; Nichols & Waschbusch, 2004).

An online CPT may be highly beneficial in terms of ecological validity, ease of administration, cost-effectiveness, and high compliance rates among participants. The opportunity to collect valid and reliable CPT assessments over the Internet opens new and exciting possibilities for clinical research and for follow-up on clinical care. However, this important potential should be viewed in light of some limitations. First, although the results of the present studies support the reliability and validity of the OCPT in young healthy adults as well as in an adult clinical sample with ADHD, the psychometric properties of the OCPT should be further examined on larger samples of populations with a wider age range. Second, the outcome of unsupervised online tests might be distorted by inappropriate testing environments or by deliberate misuse of the test. Clearly, more control over testing conditions can be achieved in the lab or the clinic. The relative lack of control at the home environment might make it difficult to determine whether observed impairments are attributable to testing conditions or to the person's inherent variables. The use of webcams along with OCPT may be recommended to capture some of the qualitative information that neuropsychologists typically capture in the lab or clinic as well as response of patients to external distractions during testing (e.g., a ringing telephone). The use of webcams should be quite easy to implement and may be useful for both clinical and research settings.

In conclusion, the preliminary results of both Study 1 and Study 2 support the reliability and validity of the online CPT as a tool for measuring key variables of attention function in young adults and therefore support its utility as a tool for the assessment of attention function in field research and clinical setups.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Note

1. In the present study, all participants took the home test first and the lab tests a week later. To alleviate, to some extent, the concern of order effects, an additional data set from 15 participants (12 females) was collected in a reversed order (i.e., Online Continuous Performance Test and Continuous Performance Tests-II in the lab first and then the Online Continuous Performance Test at home a week later). Analyses of these data in relation to the data reported in Study 1 showed no order effects and the same pattern of test–retest correlations as reported here. Exploratory analyses combining the reversed-order data with the data reported here resulted in higher levels of test–retest reliabilities.

References

Advokat, C., Martino, L., Hill, B. D., & Gouvier, W. (2007). Continuous Performance Test (CPT) of college students with ADHD, psychiatric disorders, cognitive deficits, or no diagnosis. *Journal of Attention Disorders, 10*, 253-256.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anastasi, A., & Urbina, S. (1998). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.

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

Barkley, R. A. (1998). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (2nd ed.). New York, NY: Guilford.

Bora, E., Yucel, M., & Pantelis, C. (2009). Cognitive endophenotypes of bipolar disorder: A meta-analysis of neuropsychological deficits in euthymic patients and their first-degree relatives. *Journal of Affective Disorders, 113*, 1-20.

Buchanan, R. W., Strauss, M. E., Breier, A., Kirkpetrick, B., & Carpenter, W. T. (1997). Attentional impairments in deficit and nondeficit forms of schizophrenia. *American Journal of Psychiatry, 154*, 363-370.

Conners, C. K. (2004). *Conners' Continuous Performance Test (CPT II) version 5 for Windows, technical guide and software manual*. Toronto, Ontario, Canada: Multi-Health Systems.

Conners, C. K., Epstein, J. N., Angold, A., & Klaric, J. (2003). Continuous Performance Test in a normative epidemiological sample. *Journal of Abnormal Child Psychology, 31*, 555-562.

Corkum, P. V., & Siegel, L. S. (1993). Is the continuous performance task a valuable research tool for use with children with attention deficit hyperactivity disorder? *Journal of Child Psychology and Psychiatry, 34*, 1217-1239.

DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (1998). *ADHD Rating Scale-IV: Checklists, norms and clinical interpretation*. New York, NY: Guilford.

Egeland, J., & Kovalik-Gran, I. (2010). Validity of the factor structure of Conners' CPT. *Journal of Attention Disorders, 13*, 347-357.

Epstein, J. N., Conners, C. K., Sitarenios, G., & Erhardt, D. (1998). Continuous Performance Test results of adults with attention deficit hyperactivity disorder. *Clinical Neuropsychologist, 12*, 155-168.

Epstein, J. N., Erkanli, A., Conners, C. K., Klaric, J., Costello, J. E., & Angold, A. (2003). Relations between continuous performance test, performance measures and ADHD behaviors. *Journal of Abnormal Child Psychology, 31*, 543-554.

Epstein, J. N., Johnson, D. E., Varia, I. M., & Conners, C. K. (2001). Neuropsychological assessment of response inhibition in adults with ADHD. *Journal of Clinical and Experimental Neuropsychology, 23*, 362-371.

Forbes, G. (1998). Clinical utility of the Test of Variables of Attention in the diagnosis of attention-deficit/hyperactivity disorder. *Journal of Clinical Psychology, 54*, 461-476.

Goel, N., Rao, H., Durmer, J. S., & Dinges, D. F. (2009). Neurocognitive consequences of sleep deprivation. *Seminars in Neurology, 29*, 320-339.

Greenberg, L., Kindschi, C., Dupuy, T., & Corman, C. (1996). *Test of variables of attention*. Los Alamitos, CA: Universal Attention Disorders.

Greenberg, L. M., & Waldman, I. D. (1993). Developmental normative data on the Test of Variables of Attention (TOVA). *Journal of Child and Adolescent Psychiatry, 34*, 1019-1030.

Halperin, J. M., Sharma, V., Greenblatt, E., & Schwartz, S. T. (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.

Hill, B. D., Pella, R. D., Singh, A. N., Jones, G. N., & Gouvier, W. D. (2009). The Wender Utah Rating Scale: Adult ADHD diagnostic tool or personality index? *Journal of Attention Disorders, 13*, 87-94.

Klein, C., Wendling, K., Huettner, P., Ruder, H., & Peper, M. (2006). Intra subject variability in attention-deficit hyperactivity disorder. *Biological Psychiatry, 60*, 1088-1097.

Kooij, J. J. S., Boonstra, A. M., Swinkels, S. H. N., Bekker, E. M., de Noord, I., & Buitelaar, J. K. (2008). Reliability, validity, and utility of instruments for self-report and informant report concerning symptoms of ADHD in adult patients. *Journal of Attention Disorders, 11*, 445-458.

Leark, R. A., Wallace, D. R., & Fitzgerald, R. (2004). Test-retest reliability and standard error of measurement for the Test of Variables of Attention (TOVA) with healthy school-age children. *Assessment, 11*, 285-289.

Llorente, A., Amado, A., Voigt, R., Berretta, M., Fraley, J., Jensen, C., & Heird, W. C. (2001). Internal consistency, temporal stability, and reproducibility of individual index scores of the Test of Variables of Attention in children with attention -deficit/hyperactivity disorder. *Archives of Clinical Neuropsychology, 16*, 535-546.

Llorente, A. M., Voigt, R., Jensen, C. L., Fraley, J. K., Heird, W. C., & Rennie, K. M. (2008). The Test of Variables of Attention (TOVA): Internal consistency (Q1 vs. Q2 and Q3 vs. Q4) in children with Attention Deficit/Hyperactivity Disorder (ADHD). *Child Neuropsychology, 14*, 314-322.

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.

McGee, R. A., Clark, S. E., & Symons, D. K. (2000). Does the Conners' Continuous Performance Test aid in ADHD diagnosis? *Journal of Abnormal Child Psychology, 28*, 415-424.

Nichols, S. L., & Waschbusch, D. A. (2004). A review of the validity of laboratory cognitive tasks used to assess symptoms of ADHD. *Child Psychiatry and Human Development, 34*, 297-315.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

O'Brien, J. D., Halperin, J. M., Newcorn, J. H., Shrma, V., Wolf, L., & Morganstein, A. (1992). Psychometric differentiation of conduct disorder and attention deficit disorder with hyperactivity. *Journal of Developmental and Behavioral Pediatrics, 13*, 274-277.

Pastor, P. N., & Reuben, C. A. (2008). Diagnosed attention deficit hyperactivity disorder and learning disability: United States, 2004-2006. *Vital Health Statistics, 237*, 1-14.

Rabin, L. A., Barr, W. B., & Burton, L. A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA division 40 members. *Archives of Clinical Neuropsychology, 20*, 33-65.

Riccio, C. A., & Reynolds, C. R. (2001). Continuous performance tests are sensitive to ADHD in adults but lack specificity. A review and critique for differential diagnosis. *Annals of the New York Academy of Sciences, 931*, 113-139.

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, NY: John Wiley.

Sadeh, A., Dan, O., & Bar-Haim, Y. (2011). Online-assessment of sustained attention following sleep restriction. *Sleep Medicine, 12*, 257-261.

Walker, A. J., Shores, E. A., Trollor, J. N., Lee, T., & Sachdev, P. S. (2000). Neuropsychological functioning of adults with attention deficit hyperactivity disorder. *Journal of Clinical and Experimental Neuropsychology, 22*, 115-124.