Summary of IVA-AE2 Research Studies

The original Diagnostic Validity, Concurrent Validity, and Test-Retest Reliability studies for IVA-AE are presented below. Since the IVA-AE2 test procedures and its construction do not differ from the original IVA-AE, these research studies are applicable for both versions of this test. Thus, in the studies presented below the test will be referred to as the IVA Advanced Edition CPT.

Diagnostic Validity Study of the IVA-Advanced Edition

This study was conducted to determine whether the IVA-Advanced Edition Continuous Performance Test (CPT) was able to discriminate individuals who believed that they had attention problems from a matched group of subjects who did not specify having any attentional difficulties. This study was completed at St. Cloud University. The authors of this study are Tim Tinius, Ph.D. and Joseph A. Sandford, Ph.D.

Introduction

It is important in validating tests that they be evaluated in terms of their ability to discriminate individuals who report the identified problems that the test purports to assess. In this case, individuals who reported believing that they had ADHD symptoms were selected based on their questionnaire responses. These individuals were tested using the IVA Advanced Edition

This test was specifically created to be more complex than most CPTs in order to better evaluate adults with attention problems. Based on these test results, it could be determined whether this test is valid and able to differentiate this group with reported attention problems from other individuals who do not report any attention problems whatsoever. This provides an independent measure of validity, based on the specific criterion that the test is designed to assess.

Participants

Twenty one individuals (6 males and 15 females) were selected for participation in this study. These individuals completed the Extended Version of the test. All of these individuals were selected for this study, because they reported in a questionnaire that they believed they had ADD or ADHD. They were matched against a normative group by age, sex and education. The groups of the study were selected to remove people who reported a history of brain injury, a seizure disorder, that they were currently taking medication or who were currently involved in counseling. The mean age was 21 and the age range was 18-23. All subjects were volunteers.

Procedure

The standardized procedure for administering this test, as described in its administrative manual, was followed. The instructions were displayed on the computer monitor and also spoken in a clear, digitized, female voice through headphones. All participants were given the opportunity to learn how to click the mouse correctly and a two-button USB optical mouse was used a test input device. Individuals tested sometimes alone, but sometimes these volunteers were tested on separate computers in small group format. In all cases, privacy was maintained.

The test begins with a warm-up session for both auditory and visual targets and a practice session is then completed before the main part of the test begins. The warm-up session consisted of 20 auditory targets and 20 visual targets. This was followed by the 100 item practice session, during which the individual taking the test was given immediate auditory and visual feedback as to whether they made mistakes or not (i.e., the computer said “Oops!). Whenever any participant makes 15 errors, the practice test will be briefly interrupted and the test instructions will then be repeated again by the computer. These 15 errors could be either errors of omission or commission and did not have to be made in sequential order to trigger this instructional help. The IVA-AE2 CPT Extended Test consists of 1000 stimuli presented at one per second. The test rules required participants to click whenever they saw a “3" or heard a “5". For each trial, one number was both displayed and another different number spoken. The numbers used as foils were 1, 2, 4, 5, 6, and 7, and they were displayed using different fonts and spoken using different male and female voices. The total testing time was about 20 minutes in length. Participants then completed the IVA-AE2 Self-Report Scale regarding their test experience. They were thanked and debriefed, but no feedback regarding their scores was provided.

Once the test began, there was no help provided by the test examiner, except as allowed in the respective test administration manual. So, if the subjects removed their finger from the left mouse button or switched to the right mouse button, they were asked to resume clicking only the left button. If they asked any questions, then the standard response made was “Keep working. Do your best.” Scores were automatically saved by the computer for later analysis.

Results

The results are presented in Table 1. This table shows the Global and Primary scale quotient scores for the IVA-AE2 CPT (Extended Version). Scores are presented for both the group that believes they have ADHD and randomly selected matched pairs from the normative sample. The results in this table showed that the eight Global scales all showed a significant difference. For the Response Control Auditory and Visual Global scales, the difference was significant at the $p<.05$ level and the normative group was higher by 8 and 11 points, respectively. A greater difference was seen for the Attention Quotient Auditory and Visual Global scales and the increase was 17 points higher for the normative group and this difference is more than one standard deviation. The significance for these scales was at the $p<.01$ and $p<.001$ levels, respectively. The Sustained Attention Auditory scale, Visual Global scales and Mental Concentration Auditory and Visual Global scales also showed a large difference. Thus, all of the Global scales were significantly different and the ADHD "symptom" group was identified as scoring poorer in terms of their attentional functioning than the matched pairs comparison "normal" group.

The Primary scales varied in their significance. The Comprehension scale, which has been identified in the IVA-2 CPT as one of the most sensitive indicators of ADHD symptoms, was found in the IVA-AE2 CPT to also be a very sensitive discriminator. The matched comparison group was 23 points higher for the Comprehension Auditory scale and 22 points higher for the Comprehension Visual scale. Both these scores were judged to be significant ($p<.01$ and $p<.05$, respectively). It is interesting to note that the Fine Motor Control scale also was significantly different and individuals with attention problems were 13 points lower than the matched group. The Focus Visual scale did show a significant difference at the $p<.05$ level and the matched group without attention problems was 9 points higher than the other group. The Steadiness Auditory and Visual scale, which specifically measures the subject's ability to maintain attention under high demand conditions, did show a significant difference both for the auditory and visual domains. The individuals without attention problems were 25 points higher on the auditory domain and 21 points higher on the visual domain. Lastly, the Vigilance scale, which specifically measures errors of omission, showed a significant difference with the comparison group without attention problems being 18 points higher for the Vigilance Auditory Primary scale and 21 points higher for the Vigilance Visual Primary scale. All the other scales were not significantly different.

		Quotient Score Means			
Global Scales	**Type**	**ADHD**	**Norm**	**Difference**	**Significance**
Response Control	Auditory	96	104	8	p < .05
	Visual	92	103	11	p < .05
Attention	Auditory	86	103	17	p < .01
	Visual	87	104	17	p < .001
Sustained Attention	Auditory	84	102	18	p < .001
	Visual	85	104	19	p < .01
Mental Concentration	Auditory	86	104	18	p < .001
	Visual	86	104	18	p < .01
Primary Scales					
Comprehension	Auditory	83	106	23	p < .001
	Visual	84	106	22	p < .05
Consistency	Auditory	96	101	5	n.s.
	Visual	95	100	5	n.s.
Fine Motor Control	Motoric	91	104	13	p < .05
Focus	Auditory	98	101	3	n.s.
	Visual	94	103	9	p < .05
Prudence	Auditory	96	104	8	n.s.
	Visual	96	104	8	n.s.
Reliability	Auditory	94	101	7	n.s.
	Visual	94	101	7	n.s.
Sensory/Motor	Auditory	105	100	-5	n.s.
	Visual	102	99	-3	n.s.
Speed	Auditory	93	101	8	n.s.
	Visual	95	99	4	n.s.
Stamina	Auditory	100	102	2	n.s.
	Visual	94	101	7	n.s.
Steadiness	Auditory	82	107	25	p < .001
	Visual	85	106	21	p < .05
Stillness	Motoric	88	88	0	n.s.
Vigilance	Auditory	85	103	18	p < .01
	Visual	85	106	21	p < .01

n.s. = Not Significant

Table 1: Global and Primary Scale Quotient Score Differences between a group that reported ADHD symptoms and a paired normative sample for the IVA Advanced Edition CPT (Extended Version).

Table 2 was provided in order to examine whether the IVA-AE2 CPT Basic Version, which consists of the first 500 stimuli trials was also able to differentiate the two groups. While significance was found that was very similar to the findings of the Extended Test, the differences were noticeably larger and more profound in the Extended Version. This provides evidence which supports that the Extended Version is the better discriminating they have to sustain attention, then attention problems are more easily identified in adults. Generally speaking, the same pattern of differences was found for both Global test and this makes logical sense. Adults are often able to pay attention for short periods of time, even if they have attention difficulties, when they are motivated. When they have to sustain attention, then attention problems are more easily identified in adults. Generally speaking, the same pattern of differences was found for both Global and Primary scales on the short test version. In Table 2 it is clearly shown that the shorter version of the test is also accurate and does discriminate adults who believe they have ADHD from adults who do not report having this disorder.

		Quotient Score Means			
Global Scales	**Type**	**ADHD**	**Norm**	**Difference**	**Significance**
Response Control	Auditory	96	104	8	n.s.
	Visual	96	104	8	p < .05
Attention	Auditory	89	105	16	p < .01
	Visual	89	104	15	p < .01
Sustained Attention	Auditory	87	102	15	p < .01
	Visual	90	105	15	p < .01
Mental Concentration	Auditory	88	105	17	p < .01
	Visual	90	105	15	p < .01
Primary Scales					
Comprehension	Auditory	86	107	21	p < .001
	Visual	86	106	20	p < .05
Consistency	Auditory	98	102	4	n.s.
	Visual	94	101	7	p < .05
Fine Motor Control	Motoric	92	104	12	n.s.
Focus	Auditory	98	102	4	n.s.
	Visual	95	105	10	p < .01
Prudence	Auditory	96	104	8	n.s.
	Visual	97	106	9	n.s.
Reliability	Auditory	92	100	8	n.s.
	Visual	95	100	5	n.s.
Sensory/Motor	Auditory	105	100	-5	n.s.
	Visual	102	99	-3	n.s.
Speed	Auditory	94	102	8	n.s.
	Visual	96	99	3	n.s.
Stamina	Auditory	98	102	4	n.s.
	Visual	94	101	7	n.s.
Steadiness	Auditory	86	108	22	p < .001
	Visual	87	107	20	p < .01
Stillness	Motoric	90	88	-2	n.s.
Vigilance	Auditory	88	104	16	p < .01
	Visual	91	107	16	p < .01

n.s. = Not Significant

Table 2: Global and Primary Scale Quotient Score Differences between a group that believes to have ADHD and a paired Normative sample for the IVA-AE2 CPT (Basic Version).

An analysis was also completed to identify the diagnostic accuracy of the IVA-AE2 test results using its flowchart diagnostic algorithm. As discussed above, volunteers were selected based on their report that they believed for one reason or another that they had ADD or ADHD. The study was limited in any further verification of this possible diagnosis and, thus, this group of individuals can best be described as having a "self-labeled" diagnosis. The results of this analysis are presented in Table 3a and 3b below.

Table 3a. Comparison of Clinical Diagnosis and IVA-2 Diagnosis for ADHD

		Self-Diagnosis		
		ADHD	No ADHD	Total
IVA-AE2 Test Diagnosis	ADHD	15	2	17
	No ADHD	6	19	25
	Total	21	21	42

Table 3b. Diagnostic Accuracy of the IVA-AE2 Test for Self-Diagnosed ADHD

	Diagnostic Accuracy
Sensitivity	71%
Specificity	90%
Positive Predictive Power	88%
Negative Predictive Power	76%

Sensitivity is defined as the probability of the test identifying a positive result given that the individual has ADHD. Specificity is the probability of the test accurately identifying individuals who do not have ADHD. Positive Predictive Power is the percentage of individuals diagnosed by the test as having ADHD who were also clinically diagnosed. Negative Predictive Power is the percentage of patients that were correctly diagnosed as not having ADHD.

Discussion

The results presented in Table 1 for the IVA-AE2 CPT Extended Version clearly document that attentional difficulties are discriminated by the IVA-AE2 CPT for individuals who report that they believe they have ADHD. While it is recognized that many people who report they have ADHD may demonstrate attention problems, it is possible that these problems may be due to other factors. The value of this particular study was to clearly document the validity of the test in identifying individuals who believed they had attention problems. From one point of view, it may be that individuals who report they have attention problems may be less motivated to do well on the test and not perform their best, due to what may be best described as a "self-fulfilling prophecy." In order to rule out this potential problem, a self-report questionnaire administered at the end of the test was analyzed. All individuals taking the test were asked to respond to the question "Did you try very hard to do your best on this test?" The possible responses included: Not at All, Some, A Lot or Very Much. These responses were assigned a rating from 1-4, respectively. Individuals who did report that they had ADHD symptoms had a mean rating of 3.6 and individuals who did not report problems had a mean rating of 3.2. These ratings did not differ and showed that both groups were highly motivated to perform to their very best.

It is interesting to note that another question that was asked, "Was it hard to sit still during the test?" did significantly differentiate between the two groups. In response to this question, individuals who reported ADHD symptoms had a mean score of 3.1, indicating that they had difficulty sitting still a lot or very much and individuals without any reported attention problems had a mean score of 2.0, indicating that it was somewhat hard to sit still during the test.

This questionnaire differentiated the two groups at a very significant level ($p<.01$). Also, the question "Did taking the test make you feel mad, angry or frustrated?" differentiated the two groups. Individuals who reported ADHD symptoms had a mean of 2.3, which indicated that they felt some or a lot of these negative feelings while taking the test and individuals who did not believe they had the ADHD diagnosis had a rating of 1.6, indicating that they rated the test's emotional impact between not at all and somewhat stressful. This difference was significant at the $p<.05$ level. Thus, it was found that both ADHD symptom and normal groups were equally motivated to perform well. However, the self-report ratings regarding the difficulties they experienced sitting still during the test and their negative emotions during the test were significantly different for these two groups. These self-report findings are congruent for those individuals who reported that they believed themselves to suffer from ADHD.

Overall, the Extended Test was observed to be more sensitive than the Basic Test in differentiating individuals who believed they had ADHD. It is not surprising that the Extended Test provides a more accurate measure of attentional dysfunction given that it is twice as long. Thus, the Extended Test will be recommended as the version of choice to use with the IVA-AE2 CPT in order to better and more clearly differentiate an individual's problems related to response control and attentional difficulties. It was also found that all of the Global scales did show significant difference, but the primary difference identified involved difficulties with sustaining attention throughout the test and not impulsivity. The three main Global scales, each of which separately assesses both auditory and visual modalities clearly showed a significant difference between these two groups. This supports the diagnostic validity of the IVA-AE2 CPT.

The diagnostic accuracy of the IVA-AE2 test showed that based solely on the test results it was possible to correctly identify about seven out of ten individuals who reported that they believed they had ADD or ADHD. The test results also found that the test was not likely to support a diagnosis of ADHD when an individual does not report significant ADHD or ADD type symptoms. The test findings found that only two out of twenty-one individuals (10%) from the matched control group who did not report that they were ADHD, as having attention deficits that indicated the possibility of this diagnosis. Thus, in clinical practice the IVA-AE2 will likely lead to very few false ADHD positive diagnoses.

The issue of the test's sensitivity is likely to be higher, if a formal clinical diagnosis of the individuals who reported significant ADHD type symptoms was completed given that many other problems such as anxiety, learning disabilities, depression, etc. can impact attentional functioning and mislead an individual into believing that he or she has ADHD when this may not be the case. In addition, based on the IVA-2 diagnostic study the combination of the IVA-AE2 test findings with rating scale data is likely to increase the diagnostic accuracy to about 80% for self-diagnosed individuals. Further research with the IVA-AE2 is warranted to expand on this study. These results do though support that this test is sufficiently accurate to help clinicians in making their diagnosis of ADHD or ruling it out.

Concurrent Validity Test Study of the IVA-Advanced Edition

This study was conducted to compare the correlation and quotient scale scores of the IVA-2 Continuous Performance Test (CPT) with the new IVA-Advanced Edition CPT. This study was completed at St. Cloud University. The authors of the study are Tim Tinius, Ph.D. and Joseph A. Sandford, Ph.D.

Introduction

In order for a test to be considered valid, it is often compared with another test that is recognized as a valid measure of the relevant psychological functions in order to determine concurrent validity. The IVA-AE2 CPT is a new and modified version of the IVA-2 CPT. The IVA-2 CPT was designed to be fairly simple in its construct in order that it could be used to assess a wide range of ages. The IVA-AE2 was specifically created to be more complex in order to better evaluate adults with attention problems. Thus, the IVA-2 can be used with the age ranges of 6 to 96 and the

IVA-AE2 is limited to the age range of most adults in their prime, which are the ages of 18 to 50. By comparing these two tests, it can be determined whether individuals taking the IVA-AE2 CPT score similarly to individuals taking the IVA-2 CPT. Numerous research studies have clearly established the validity and clinical usefulness of the IVA-2 test. Thus, if the IVA-AE2 CPT is valid, it would be expected that its scale scores would be highly correlated with the IVA-2 CPT scales. Also, concurrent validity would be established if there were no significant differences for the scales in terms of their quotient scores. In this way, the IVA-AE2 CPT would be shown to be similar in its measurement of attentional functioning of adults when compared to the IVA-2 CPT.

Participants

Twelve participants were selected for this study. Participants first completed the IVA-2 CPT and then completed the IVA-AE2 CPT. There were four males in this study and eight females. Eleven of the twelve participants had some college education. No screening was done for any of the participants. One individual did report a history of a head injury and all other participants did not report any problems. The mean age was 21 and the age range was 18-23. All subjects were volunteers.

Procedure

The standardized procedure for administering the IVA-2 CPT and the IVA-AE2 CPT, as described in their respective administration manuals, was followed. 75% of the participants took the IVA-AE2 CPT before they took the IVA-2 CPT and 25% of the participants took the IVA-2 CPT first. The instructions for both tests were displayed on the computer monitor and also spoken in a clear, digitized female voice through headphones. All participants were given the opportunity to learn how to click the mouse correctly and a two-button Microsoft USB optical mouse was used as the test input device. Individuals were tested sometimes alone, but were also tested in small group format. In all cases, privacy was maintained.

The IVA-2 test begins with separate warm-up sessions for both visual and auditory targets. The warm-up for this test consists of 10 auditory and 10 visual targets. Next, the subjects were given a brief practice session of 32 test stimuli and required to click the mouse whenever they saw or heard the number “1" and not click if they saw or heard the number “2".

The IVA-AE2 CPT begins with a warm-up session for both auditory and visual targets and a practice session is then completed before the main part of the test begins. The warm-up session consisted of 20 auditory targets and 20 visual targets. This was followed by the 100 item practice session, during which the individual taking the test was given immediate auditory and visual feedback as to whether they made mistakes or not (i.e., the computer said “Oops!). Whenever any participant made 15 errors, the practice test was briefly interrupted and the test instructions were briefly repeated to them again by the computer. These 15 errors could be either errors of omission or commission and did not have to be made in sequential order to trigger this instructional help. The basic version of the test was administered, but only the Basic Test was used in this study. The total testing time was about 20 minutes in length. Participants then completed the IVA-AE2 Self-Report Scale regarding their test experience. They were thanked and debriefed, but no feedback regarding their scores was provided for either test.

The IVA-AE2 CPT main test (for the Basic Test Version analysis) consists of 500 stimuli presented at one per second. The test rules required participants to click whenever they saw a “3" or heard a “5". For each trial, one number was both displayed and another different number spoken. The numbers used as foils were 1, 2, 4, 5, 6, and 7, and they were displayed using different fonts and spoken using different male and female voices. Thus, the IVA-AE2 CPT required much faster and more complex attentional processing than the IVA-2 CPT.

Once either test began, there was no help provided by the test examiner, except as allowed in the respective test administration manual, and these rules were the same for both tests. So, if the subjects removed their finger from the left mouse button or switched to the right mouse button, they were asked to resume clicking only the left button. If they asked any questions, then the standard response made was "Keep working. Do your best." Scores were automatically saved by the computer for later analysis.

Results

The results are presented in Table 1. This table shows the Global and Primary scale quotient score correlation of the IVA-AE2 CPT (Basic Version) with the IVA-2 CPT. The Basic Version of the IVA-AE2 CPT consists of the same number of trials as the IVA-2 CPT, which is 500. For the Global scales, 88% of them showed a significant correlation. The correlations for the Global scales ranged from a low of -.05 for the Attention Auditory Global scale to a high of .69 for the Attention Visual Global scale. The Global test scale scores were found for the Response Control Auditory and Visual scales to be correlated in the moderate range. The Attention Auditory scale was not significantly correlated, but the Attention Visual scale was found to be correlated in the marked range with a correlation of .69. The Sustained Attention Auditory Global scale was found to be mildly correlated and the Sustained Attention Visual Global scale was found to be markedly correlated. The Mental Concentration Auditory Global scale was found to be mildly correlated and the Mental Concentration Visual Global scale was found to be marked in its correlation. Thus, almost all of the Global scales did show a significant correlation between the two tests.

The Primary scales had correlations that ranged from a low of -.07 for the Reliability Visual scale to a high of .72 for the Prudence Auditory Scale and the Steadiness Visual scale. 81% of the Primary scales did show a significant correlation and many of these correlations were in moderate to marked range. There were four scales that did not show a significant correlation between the two tests. These scales were the Fine Motor Control scale, the Reliability Visual scale, Speed Auditory scale and Vigilance Auditory scale. Thus, the Primary scales, which are based on the raw scores, did show a very high concordance between the two tests.

Global Scales	Type	Correlation	Strength
Response Control	Auditory	0.49	2
	Visual	0.41	2
Attention	Auditory	-0.05	n.s.
	Visual	0.69	3
Sustained Attention	Auditory	0.28	1
	Visual	0.65	3
Mental Concentration	Auditory	0.24	1
	Visual	0.60	2
Primary Scales			
Comprehension	Auditory	0.58	2
	Visual	0.69	3
Consistency	Auditory	0.44	2
	Visual	0.51	2
Fine Motor Control	Motoric	-0.16	n.s.
Focus	Auditory	0.47	2
	Visual	0.62	3
Prudence	Auditory	0.72	3
	Visual	0.52	2
Reliability	Auditory	0.72	3
	Visual	-0.07	n.s.
Sensory/Motor	Auditory	0.69	3
	Visual	0.62	3
Speed	Auditory	0.04	n.s.
	Visual	0.63	3
Stamina	Auditory	0.35	1
	Visual	0.39	1
Steadiness	Auditory	0.31	1
	Visual	0.72	3
Vigilance	Auditory	-0.13	n.s.
	Visual	0.60	2
Strength Legend	**Values**		
n.s. = Not Significant	<.20		
1 Mild Correlation	.20-.39		
2 Moderate Correlation	.40-.59		
3 Marked Correlation	.60-.79		
4 Very High Correlation	.80-1.00		

Table 1: Global and Primary Scale Quotient Score Correlation of the IVA-AE2 CPT (Basic Version) and the IVA-2 Test Administrations.

An examination was also made of the quotient scale scores to identify if there were any significant differences in Global and Primary scale scores of the IVA-2 CPT as compared to the IVA-AE2 CPT. The results of this analysis are presented in Table 2 of this study. A paired sample, two-tailed t-test was used to identify any significant differences. The significance cutoff level for this study was set at $p<.05$. No significant differences were found for any of the global scales of these two tests. For the Primary scales, there were six scales that did show a significant difference. A quick review of these Primary scales does show some variability with both higher and lower scale scores between the two tests. Overall, it was found that 71% of the Primary scale scores did not significantly differ.

The examination of the Primary scale scores did identify some differences. The Focus Auditory scale was significantly higher ($p<.05$) and increased a mean of 14 points from the subjects' scores on the IVA-2 CPT to their scores on the IVA-AE2 CPT. The Prudence Visual scale score showed a 13 point increase for the IVA-AE2 CPT which was significant ($p<.05$). The Sensory/Motor Visual scale of the IVA-AE2 CPT was 11 points less, and this decrease was significant ($p<.01$). The Speed Visual scale score was 16 points lower on the IVA-AE2 CPT, and this score was significant ($p<.001$). Also, the Stamina Auditory scale was found to be 13 points higher for the IVA-AE2 CPT which was significant ($p<.05$). Finally, the Vigilance Visual scale was significantly higher by 19 points ($p<.05$) for the IVA-AE2 CPT.

Discussion

An examination of the results of this concurrent validity study, found only a few significant differences between the two tests which primarily involved auditory stimuli. Participants taking the IVA-AE2 CPT were found to respond to auditory stimuli in a slightly different way due to the complexity and challenging nature of this test. As noted above, in the description of the test, the IVA-2 CPT is very simple and the individual has to remember only one rule, which is to click when they see or hear the number "1". The foil in the IVA-2 CPT test is only one other number and that is the number, two. In addition, only one test stimuli, either a visual or auditory stimuli, is presented at a time. In contrast, the IVA-AE2 CPT has two different targets (i.e. a visual "3" and an auditory "5") and both visual and auditory stimuli are presented simultaneously for each test trial. For the IVA-AE2 CPT, each trial consists of both a target and a non-target. The visual stimuli are also much more varied in their shape, because different fonts are used for the various numbers. Different voices, which include both male and female, are also used for the spoken numbers in the IVA-AE2 CPT. Thus, the IVA-AE2 CPT is a much more challenging test that requires more mental discrimination, especially for the auditory stimuli, because different voices are used to say the auditory target. In contrast, the visual targets are more similar and consistent. While the three is displayed in

		Quotient Score Means			
Global Scales	**Type**	**IVA-2**	**IVA-AE**	**Difference**	**Significance**
Response Control	Auditory	85	98	13	n.s.
	Visual	89	101	12	n.s.
Attention	Auditory	86	99	13	n.s.
	Visual	91	96	5	n.s.
Sustained Attention	Auditory	85	94	9	n.s.
	Visual	83	94	11	n.s.
Mental Concentration	Auditory	85	95	10	n.s.
	Visual	87	95	8	n.s.
Primary Scales					
Comprehension	Auditory	94	87	-7	n.s.
	Visual	97	88	-9	n.s.
Consistency	Auditory	89	96	7	n.s.
	Visual	94	99	5	n.s.
Fine Motor Control	Motoric	102	97	-5	n.s.
Focus	Auditory	83	97	14	p < .05
	Visual	94	99	5	n.s.
Prudence	Auditory	88	96	8	n.s.
	Visual	90	103	13	p < .05
Reliability	Auditory	90	88	-2	n.s.
	Visual	95	96	1	n.s.
Sensory/Motor	Auditory	110	104	-6	n.s.
	Visual	113	102	-11	p < .01
Speed	Auditory	108	100	-8	n.s.
	Visual	108	92	-16	p < .001
Stamina	Auditory	91	104	13	p < .05
	Visual	97	99	2	n.s.
Steadiness	Auditory	96	89	-7	n.s.
	Visual	92	88	-4	n.s.
Vigilance	Auditory	83	101	18	n.s.
	Visual	81	100	19	p < .05
n.s. = Not Significant	<.20				

Table 2: Global and Primary Scale Quotient Score Differences for the IVA-AE2 CPT (Basic Version) and the IVA-2 Test Administrations.

different fonts, all the numbers still look like a three, and the fonts are not distorted in any way. It is important to note though, that based on Table 2 that the Attention Auditory Global scale does not significantly differ between the two tests, and so the normative data helps to correct for the complexity and differences of the auditory stimuli in the IVA-

AE2 CPT. A very high concurrent validity was established for the Global scales for the IVA-AE2 CPT, because no significant difference was found between its scale scores and those of the IVA-2 CPT.

While the Fine Motor Control scale was not found to be correlated, it did not significantly differ for the two tests. The Reliability Visual scale also did not show a significant correlation, and its quotient scores did not differ between the two tests. Likewise, the Speed Auditory scale did not show a significant difference, even though it was not correlated. Finally, even though the Vigilance Auditory scale was not found to have a significant correlation, there was no significant difference between the quotient scores of the two different tests. Thus, even when the correlation data suggested a difference in the way participants processed each test, in no case were these differences significant in respect to the scale scores.

In summary, the majority of the IVA-AE2 and IVA-2 CPT scales are significantly correlated. 88% of the Global scale scores and 81% of Primary scales were found to be significantly correlated. None of the Global scales show any significant difference between the two tests. Thus, test-takers would be expected to score similarly on both tests in respect to response control and the various measures of attentional functioning. The Primary scale scores did not show any significant difference 71% of the time. In the analysis of the test, much greater weight is placed on the Global scales which are more reliable for both tests and the variability of the Primary scales was considered as likely to be reflective of the differences in the tests due to the degree of complexity and test stimuli. Overall, the findings of this study showed that the IVA-AE2 CPT demonstrated a very high concurrent validity with the IVA-2 CPT.

Test-Retest Reliability Study

Tim Tinius, Ph.D. and Joseph A. Sandford, Ph.D. completed a test-retest reliability study of the IVA-Advanced Edition Continuous Performance Test (CPT). This study was completed at St. Cloud University in Minnesota.

Introduction

Reliability refers to the consistency of test score performance for repeated testing of an individual under similar conditions. This type of reliability test provides an index about the stability of the IVA-AE2 CPT scores over time (Anastasi, 1988) that allows therapists to evaluate whether any changes observed in the test scale scores reflect differences in a person's performance that are not solely due to either practice effects or random errors. A test that is reliable will show moderate, marked or very high correlations of its standardized scale scores for the same individuals who take it on different days. Also, reliability can be measured in respect to the consistency of test scale scores taken on two different occasions by the same people. Thus, tests that are reliable will generally result in correlated and consistent test scores for the same person over time. This study was conducted in order to demonstrate the reliability and consistency of the IVA-AE2 CPT scores over time.

In order to be used clinically in diagnostic decision-making, tests need to be evaluated for reliability. Also, the use of a test instrument for the evaluation of medication and/or treatment effects needs to take into account potential practice effects. This study addressed the test-retest reliability of the major IVA-AE2 CPT scales. The IVA-AE2 CPT has both Global and Primary scales. The difference between Global and Primary scales is that the Global scales are based on Primary scale quotient scores and that the Primary scales are based on raw data. Since the Global scale scores are based on multiple Primary scales, they have generally been found to be more stable and consistent measures of test performance than individual Primary Scales. Thus, the Global scales are the most relevant ones to examine in evaluating the IVA-AE2 CPT's reliability. The Global scaled scores examined included Response Control, Attention, Sustained Attention and Mental Control. Quotient scaled scores were compared in the analysis using the IVA-AE2 CPT normative database created on January 20, 2007.

Participants

The individuals that participated were screened for possible problems and factors that could impair test performance. Individuals reporting a history of traumatic brain injury, seizures or taking any prescribed medication for ADHD were excluded. If the individuals reported taking any other type of medication other than birth control (N=15), then their test results were included, but only if they reported taking the same medication on both

test occasions. Generally, the second test was administered about a week after the first test, and this tended to minimize any possible changes in medication. Also, if individuals were in counseling for psychological problems (N=3), they were included only if they still reported being in counseling during the time of the second test administration.

A total of 110 individuals participated in this test-retest reliability study. The test sample consisted of 21 males (19%) and 89 females (81%). The mean age was 21. The participants' ages ranged from 18 to 44. 97% of these subjects were high school graduates and 79% had some college experience. Six of the individuals who participated in the study were college graduates (5.5%). Five individuals in this study reported that they had a reading learning disability and four stated they had a math learning disability. Six individuals said that they believed they had ADHD, but did not report taking any medications for it and, thus, were included in this study. All subjects were volunteers.

Design

The standardized procedure for administering the IVA-AE2 CPT, as described in its administration manual, was followed. The second test was typically administered about one week after the participant took the first test (Mode = 7 days, Mean = 25 days). The IVA-AE2 CPT instructions were displayed on the computer monitor and also spoken in a clear, digitized female voice through headphones. Individuals were tested sometimes alone, but were also tested in small group format. In all cases, privacy was maintained.

The IVA-AE2 CPT begins with a warm-up session for both auditory and visual targets and a practice session then follows before the main part of the test begins. All participants were given the opportunity to learn how to click the mouse correctly and a two-button Microsoft USB optical mouse was used as the test input device. The warm-up session consisted of 20 auditory targets and 20 visual targets. This was followed by the 100 item practice session, during which the individual taking the test was given immediate auditory and visual feedback as to whether they made mistakes or not (i.e., the computer said "oops!). Whenever any participant makes15 errors, the practice test will be briefly interrupted and the test instructions will then be repeated again by the computer. These 15 errors could be either errors of omission or commission and did not have to be made in sequential order to trigger this instructional help. The extended version of the test was administered and total testing time was about 20 minutes in length. Participants then completed the IVA-AE2 Self-Report Scale regarding their test experience. They were thanked and debriefed, but no feedback regarding their scores was provided.

Once the test began, there was no help provided by the test examiner, except as allowed in the IVA-AE2 CPT Administration manual. So, if the subjects removed their finger from the left mouse button or switched to the right mouse button, they were asked to resume clicking only the left button. If they asked any questions, then the standard response made was “Keep working. Do your best.” Scores were automatically saved by the computer for later analysis.

Results

The test analysis is presented in three tables below. Table 1 is a correlation of the quotient scores for the test-retest comparison. The Pearson Correlation Coefficient was computed in creating this test-retest reliability table. This correlation table shows significant correlations for the Global Response Control Scale quotient scores ranging from a low of .53 (moderate strength) for the Response Control Visual Global scale to a high of .75 (marked strength) for both the Global Mental Concentration Auditory and Visual scales. For the Primary standard scales, the correlation ranged from a low of .03 (not significant) for the Stamina Visual scale to a high of .84 (very high strength) for the Speed Visual scale. The correlations of all Global and Primary scale scores were significant (i.e., had mild to very high correlations), except for the Stamina Auditory and Visual scales. A cutoff of 15 quotient points (1 standard deviation) was used to classify any significant increases and decreases for individuals in order to understand the Stamina scale results better. In respect to the Stamina scale on the second test, six percent of the individuals clinically improved (i.e., quotient scores increased by at least one standard deviation) in their auditory stamina and twenty-six percent had clinically lower quotient scaled scores (i.e., decreased by at least one standard deviation). For the Visual Stamina scale, eleven percent improved significantly and twenty-eight percent made clinically lower scores. Thus, the Stamina scales were the only scales that were found to not be reliable. The majority of the Global and Primary scales (over 69%) showed either a marked or very high correlation (.60 or higher). Thus, almost all IVA-AE2 CPT scales were found to be reliable based on the correlation data.

Global Scales	Type	Correlation	Strength
Response Control	Auditory	0.69	3
	Visual	0.53	2
Attention Quotient	Auditory	0.73	3
	Visual	0.75	3
Sustained Attention	Auditory	0.71	3
	Visual	0.70	3
Mental Concentration	Auditory	0.75	3
	Visual	0.75	3
Primary Scales			
Comprehension	Auditory	0.62	3
	Visual	0.52	2
Consistency	Auditory	0.75	3
	Visual	0.74	3
Fine Motor Control Quotient	Motoric	0.55	2
Focus	Auditory	0.80	4
	Visual	0.71	3
Prudence	Auditory	0.73	3
	Visual	0.65	3
Reliability	Auditory	0.65	3
	Visual	0.33	1
Sensory/Motor	Auditory	0.68	3
	Visual	0.67	3
Speed	Auditory	0.74	3
	Visual	0.84	4
Stamina	Auditory	0.19	n.s.
	Visual	0.03	n.s.
Steadiness	Auditory	0.53	2
	Visual	0.48	2
Stillness	Motoric	0.36	1
Vigilance	Auditory	0.58	2
	Visual	0.62	3

Strength Legend	**Values**
n.s. Not Significant	<.20
1 Mild Correlation	.20-.39
2 Moderate Correlation	.40-.59
3 Marked Correlation	.60-.79
4 Very High Correlation	.80-1.00

Table 1: Global and Primary Scale Quotient Score Correlation of the IVA-AE2 CPT (Extended Version) for the First and Second Test Administrations.

Table 2 provides the test-retest quotient scaled score differences that were found between the first and second test. All of the Global scales showed a significant improvement in test performance; reflective of a practice effect. The increase in the quotient scores ranged from 4 to 18 quotient points. All of these improvements were significant at the $p<.001$ level, based on a paired sample t-test analysis. The Primary Standard scaled scores also showed significant improvement, except for the Stamina scales which showed a significant decline on the second test. The Stamina Auditory quotient scale score declined 8 points from the first to the second test and this was significant at the $p<.001$ level. The Stamina Visual quotient scale score decreased 5 points and this was significant at the $p<.01$ level. The Stamina scales are computed by comparing the mean reaction time of the first 200 trials in the IVA-AE2 CPT Extended Test to the mean reaction time of the last 200 trials. Thus, individuals in the second administration of the IVA-AE2 CPT Extended test did show slower discriminatory reaction times during the course of the test. This decline indicates that some physical and/or mental fatigue occurred for many participants the second time they took the test. There was no significant change in the Sensory/Motor Auditory scale. The Sensory/Motor Visual scale did show significant changes at the $p<.01$ level, but the increase was only four quotient points. Thus, the simple reaction time of participants did not differ very much from the first to the second test administration. A number of the Primary scale scores did change greater than 10 quotient points, but none showed a significant change greater than 15 quotient points, which is considered as clinically relevant. For the Global scale scores, both the Attention Auditory and Visual scales showed a significant clinical increase (greater than or equal to 15 quotient points). Also, the Mental Concentration Auditory scale showed a similar significant clinical change. The Response Control Auditory and Visual Scaled scores, while they increased significantly, showed relatively small changes of four and five quotient points, respectively.

		Quotient Score Means			
Global Scales	**Type**	**Test 1**	**Test 2**	**Difference**	**Significance**
Response Control	Auditory	99	103	4	p < .001
	Visual	100	105	5	p < .001
Attention Quotient	Auditory	97	115	18	p < .001
	Visual	97	112	15	p < .001
Sustained Attention	Auditory	98	112	14	p < .001
	Visual	96	109	13	p < .001
Mental Concentration	Auditory	97	114	17	p < .001
	Visual	97	111	14	p < .001
Primary Scales					
Comprehension	Auditory	96	110	14	p < .001
	Visual	96	106	10	p < .001
Consistency	Auditory	99	106	7	p < .001
	Visual	100	109	9	p < .001
Fine Motor Control Quotient	Motoric	98	107	9	p < .001
Focus	Auditory	100	105	5	p < .001
	Visual	99	110	11	p < .001
Prudence	Auditory	99	107	8	p < .001
	Visual	99	107	8	p < .001
Reliability	Auditory	98	102	4	p < .01
	Visual	98	103	5	p < .01
Sensory/Motor	Auditory	100	99	-1	n.s.
	Visual	100	104	4	p < .001
Speed	Auditory	98	112	14	p < .001
	Visual	99	106	7	p < .001
Stamina	Auditory	100	92	-8	p < .001
	Visual	100	95	-5	p < .01
Steadiness	Auditory	97	110	13	p < .001
	Visual	96	105	9	p < .001
Stillness	Motoric	100	92	-8	p < .001
Vigilance	Auditory	97	109	12	p < .001
	Visual	96	109	13	p < .001

n.s. = Not Significant

Table 2: Global and Primary Scale Quotient Score Differences for the IVA-AE2 CPT (Extended Version) for the First and Second Test Administration.

Table 3 is provided in order to identify how the IVA-AE2 CPT test scale scores changed on the second test administration. As noted in the interpretation manual, the IVA-AE2 CPT Extended Test version consists of two parts. Each part consists for 500 trials each. Thus, it is possible in Table 3 to look at the second test administration in order to see if individuals showed any improvement or decline from the first half to the second half of that particular test. This analysis was completed in order to determine whether or not participants would be likely to perform better on a third administration of the test.

The results showed that there was a significant decline for almost all the Global and Primary scale scores from Part 1 to Part 2. However, there were three scales that did not show significant change in either direction. The Reliability Auditory Primary scale showed a small decline of three quotient points, but this was not significant. This Reliability scale is a measure of idiopathic errors of commission, which are made under low demand conditions. In examining the Stamina Auditory and Visual scales by comparing Part 1 to Part 2, it is clear that there is no significant difference between these two parts. Given the significant decline in the majority of the Global and Primary scales from Part 1 to Part 2 for the second test administration it is clear that individuals taking the test for the second time did much better in Part 1 than Part 2.

Discussion

The test-retest reliability correlation coefficients demonstrated that the IVA-AE2 CPT is highly stable across time. The participants' scale scores on the first test were significantly correlated with the second test for almost all of the scales. Thus, these high correlation coefficients showed the IVA-AE2 CPT to be very stable over time for individuals. However, the quotient score mean differences do show that for the various Global Attention scales, there is likely to be a significant increase. The likely reason for this change is that there is some practice effect that did occur for the individuals in this sample. Thus, on the second test administration, particularly in respect to the Global Attention scales, there is likely to be a significant clinical improvement of about 15 quotient score points. One of the main contributory factors that is likely to account for the observed improvements in this study is the challenging difficulty level of this test. Generally speaking, individuals in the normative sample missed between 10% and 12% of the targets (i.e., errors of omission) and also responded incorrectly to 10-12% of the non-targets (i.e., errors of commission). Thus, in respect to the Vigilance and Prudence Primary scales the observed improvement that was found in the second test administration was reflective of a 6% to 9% mean increase in the accuracy of the test takers.

		Quotient Score Means			
Global Scales	**Type**	**Part 1**	**Part 2**	**Difference**	**Significance**
Response Control	Auditory	106	102	-4	p < .001
	Visual	110	103	-7	p < .001
Attention Quotient	Auditory	118	110	-8	p < .001
	Visual	115	107	-8	p < .001
Sustained Attention	Auditory	115	108	-7	p < .001
	Visual	113	103	-10	p < .001
Mental Concentration	Auditory	117	109	-8	p < .001
	Visual	115	105	-10	p < .001
Primary Scales					
Comprehension	Auditory	113	104	-9	p < .001
	Visual	110	100	-10	p < .001
Consistency	Auditory	108	103	-5	p < .001
	Visual	113	104	-9	p < .001
Fine Motor Control Quotient	Motoric	110	102	-8	p < .001
Focus	Auditory	106	102	-4	p < .001
	Visual	114	104	-10	p < .001
Prudence	Auditory	108	105	-3	p < .001
	Visual	109	103	-6	p < .001
Reliability	Auditory	103	100	-3	n.s.
	Visual	106	99	-7	p < .001
Speed	Auditory	115	108	-7	p < .001
	Visual	108	104	-4	p < .001
Stamina	Auditory	96	95	-1	n.s.
	Visual	97	98	1	n.s.
Steadiness	Auditory	113	105	-8	p < .001
	Visual	109	100	-9	p < .001
Vigilance	Auditory	111	106	-5	p < .001
	Visual	111	104	-7	p < .001

n.s. Not Significant

Table 3: Global and Primary Scale Quotient Score Differences for the IVA-AE2 CPT (Extended Version) for the First and Second Half of the Second Test Administration.

In examining test score improvements, one way to approach the issue is to evaluate whether the changes observed resulted in clinically different labels for the relevant global scales. This analysis found that the Attention Auditory and Visual Global scales shifted from the Average range to the Above Average range. No change in the clinical label was made for the Response Control Auditory and Visual Global scales. In respect to Sustained Attention Global scales, only the auditory one was found to shift in terms of

the descriptive ranges. In this case, an improvement was observed from the Average range to the Above Average range. The Mental Concentration Global scales also showed a significant shift in the descriptive labels for both auditory and visual stimuli. Both of the Mental Concentration Auditory and Visual Global scales advanced from the Average to the Above Average range. In looking at the primary scales and applying the same approach, the Comprehension Auditory scale was found to shift from the Average to the Above Average range. A significant shift in the Focus Visual scale was also noted. It advanced from the Average to the Above Average range. The Speed Auditory scale was another scale that was found to increase from the Average to the Above Average range, as well. Finally, the Steadiness Auditory scale was also observed to shift from the Average to the Above Average range. While almost all the other scales did show a significant difference, there was no observed change in terms of the clinical label assigned. Thus, for the majority of the primary scales (82%), it would not be expected that a person taking the test for the second time would be clinically labeled as different in respect to their performance. In the case of the other three global measures of attentional functioning, a change in the descriptive label was found to be likely for the individuals in these studies who had no reported problems or who were classified as stable in respect to their treatment and/or medications.

Even though there was a significant improvement in the quotient scores for this test from the first administration to the second administration, the analysis presented in Table 3 clearly shows that this increase occurs predominately in the first half of the second test administration. Individuals that typically do not have any psychological or emotional problems were able to improve their performance on the global scales of Attention, Sustained Attention and Mental Concentration. These improvements though, do appear to “peak” in the first half of the second test administration. Thus, it does not appear likely that on a third test administration for these individuals would make further test scale score increases.

Given that most of the individuals taking the test were well educated and were generally healthy without any disorders or who had addressed their problems through treatment or medication, it leaves open the question as to whether individuals with psychological disorders are likely to show either improvements or declines on the second test. For these individuals, the test may be more challenging if they had to take it again. Generally, the response of almost all adults taking the test was that IVA-AE2 CPT is very demanding and somewhat stressful. Further research needs to be conducted with various clinical diagnostic groups in respect to a possible practice effect for the IVA-AE2 CPT.

It also needs to be noted that the IVA-AE2 CPT has a much higher ceiling in terms of possible quotient scores than the IVA-2 CPT. In other words, most adults will have a much greater possibility to have their quotient scale scores fall in the Above Average to

Very Superior ranges with the IVA-AE2 CPT. This is due to the fact that almost all adults in the normative sample for the IVA-2 CPT made very few errors given the simplicity of the test. This fact does not impair the IVA-2 CPT in many cases from identifying individuals with attention problems, but the IVA-2 CPT is limited, to some degree, in showing higher levels of performance for “average” individuals. While clinical populations with attentional dysfunction generally have significantly lower scores on the IVA-2 CPT, there is more applicability for the IVA-AE2 CPT to be used in evaluating and identifying improvements in the Above Average and higher ranges of attentional functioning. Thus, the IVA-AE2 CPT will likely be more effective than the IVA-2 CPT for evaluating treatment improvements for people in the Low Average to Average ranges of attentional abilities.

In summary, the correlation analysis shows that individuals were relatively similar in their performance from the first to the second test. Overall, about 50% of the variance in the second test is accounted for by the first test and, thus, the correlation between the majority of the global and primary scales generally fell in the marked range or very high range. The mean quotient scale differences on the test-retest analysis did find that there is likely to be a practice effect for this test sample of “stable” individuals. Whether this stability will be found in populations with significant disorders involving attentional problems will need to be determined in a future study, Given the possibility of a practice effect, it is recommended, in using this test for the purpose of evaluating medication and/or treatment, that the clinician or researcher administer a second test within one to fourteen days and use the second test as a measure of the person’s “best performance” for evaluating further changes. In research studies, a control group can also be used to control for any potential practice effects, but again, it is recommended that a second test be administered in order to reduce the variability of a subject’s test performance due to the practice effects alone. The IVA-AE2 CPT does show a very high degree of reliability and consistency in respect to test scores from the first to the second test administration and can be used with confidence by clinicians in helping them to make diagnostic decisions.

