IVA-2 Research Studies

The IVA+Plus and IVA-2 test procedures, structure and quotient scales are the same. The major differences of the IVA-2 in comparison to the IVA+Plus are the addition of new improved interpretative reports (including a clinical report writer), integration of ADHD rating scale data in the interpretative analysis, enhanced more modern user interface, improved malingering analysis and interpretation, built-in report word processor, researcher tool kit, and an improved working diagnosis flowchart. The new features and additions to the test were made in order to make the test fully compatible with the changes documented in DSM-V. Since the test itself, including the norms, have not changed, all of the original reliability and validity research studies for the IVA+Plus apply to the IVA-2. In the discussion below of the relevant research this instrument will be referred to as the IVA CPT. Further studies that document the effectiveness and value of the IVA CPT in research are available at www.braintrain.com.

IVA-2 Test Construct Validity

The IVA-2 has a construct validity based on the well establish Continuous Performance Test (CPT) model. This model has been extensively researched and when fully implemented consists of both high and a low demand components. In high demand conditions the test taker is required to rapidly respond to targets that are frequently presented with occasional non-targets. The primary purpose of the high demand test phase is to measure impulsivity (i.e., clicking to the non-target proceeded by a series of targets). In contrast, the low demand condition is incorporated in order to assess inattention. In the low demand condition the non-targets are frequent and targets rarely presented; requiring the test taker to sustain attention throughout this inherently boring task. Both high and low demand conditions are included in the IVA-2 test construction and they alternate every 50 trials. In addition, the IVA-2 is the only CPT that integrates both visual and auditory test stimuli and, thus, does not require the administration of two separate sensory modality tests. Whereas, many individuals taking a CPT can learn to expect after a number of targets of the same sensory modality are presented that soon there will be a non-target and vice versa for the low demand condition, it is extremely challenging to apply this type of test taking strategy to the IVA-2 when presented with inter-mixed visual and auditory test stimuli. This unique construct design of the IVA-2 forces each person to pay close attention to each test item and any wavering of attention is very likely to result in either an error of impulsivity or inattention. This unique test construction design contributes to the high level of sensitivity and specificity found in the research studies discussed below.

Neuropsychological, QEEG and fMRI IVA CPT Validation with Adults

There are three significant studies that have been published that support the diagnostic validity of the IVA CPT based on published neuropsychological, QEEG and fMRI research studies. First, a study by Tinius (2003) found that adults with both mTBI and ADHD scored significantly lower on the global measures of attention and response accuracy than a control group without any identified impairments. These individuals showed specific impairments of reaction time, inattention, impulsivity, and variable in their response times to test targets. The author concluded that the IVA CPT was found to validly measure both attentional and response control impairments, because it did significantly differentiate two clinical adult groups with known cognitive deficits based on a comprehensive neuropsychological evaluation from a non-clinical "normal" population.

The second study by White et al., (2005) showed that the IVA CPT is significantly correlated with QEEG measures that are known to be diagnostic indicators of ADHD. This indicator is called the Theta/Beta ratio and it is an identifier used in an FDA approved EEG test for ADHD. The IVA CPT's Full Scale Attention quotient was correlated .78 ($p<.007$) with the Theta/Beta ratio of individuals diagnosed as having ADHD based on a comprehensive psychological evaluation. In addition, this study did not find that the Pace Serial Addition Test and the Wisconsin Card Sort tests significantly differentiated ADHD and control groups. The findings of this study provided QEEG validation of the IVA CPT and to date no other CPT has been validated using this diagnostic methodology.

The third study provided fMRI validation of the IVA CPT (Ghaziri et al., 2013). This study found significant positive correlations of fMRI measurements of white matter, ranging from .66 to .68, in brain regions associated with sustained attention and the IVA CPT measure of visual attention. In addition, individuals receiving neurofeedback significantly enhanced their IVA CPT global Full Scale attention in comparison to a sham neurofeedback control group ($p<.005$).

In conclusion, the IVA CPT has been independently validated for adults with neuropsychological assessments, QEEG measurements of impaired attention, and fMRI brain regions involved in the ability to develop sustained attention. These studies provide strong evidence of the clinical sensitivity and validity of the IVA CPT for measuring attentional functioning, because of its significant correlation with the underlying psychophysiological inter-neuronal connections in the brain's network involving attention and response control abilities.

Diagnostic and Concurrent IVA CPT Validity Study with Children

The clinical usefulness and diagnostic validity of computerized visual continuous performance tests (CPT) in the assessment and diagnosis of Attention Deficit-Hyperactivity Disorder (ADHD) has been called into question in presentations by two leading researchers (Goldman, 1994 and Halperin, 1994). Both of these researchers had concluded that the diagnostic validity of current visual computerized CPTs was not sensitive or specific to the degree that this type of test should normally be included in a multi-method assessment battery for ADHD. Barkley (1990) advanced his view that the potential for the computerized CPT was great, because this part of a comprehensive assessment was not tainted by the personal opinion biases that can occur in subjective rating scales. The problem in using a CPT arises in being able to accurately interpreting its findings at the individual level. Barkley (1994a) in a reanalysis of earlier published data reported that a visual CPT correctly classified over 90% of the children with an abnormal score, but had false negative rates of 37% or higher. In other words, a visual CPT failed to identify about 2 of 5 children previously diagnosed using other assessment techniques as ADHD.

To date, most computerized CPTs are visual, leaving out the assessment of possible auditory impulsivity and inattention problems associated with ADHD. There has been an assumption that there are no significant differences between auditory and visual CPTs, but research by Taylor (1994) found that "normal" college students were significantly more impulsive on auditory as compared to visual CPT tasks. The concurrent and diagnostic validity of the IVA CPT, which includes both auditory and visual measures of impulsivity and inattention, was explored in this study (Fine, Goldman, & Sandford, 1995). The purpose of this research was to determine whether a diagnostic classification based on the IVA CPT test results was sufficiently accurate that it would be clinically valid to use routinely for individual assessments.

The first group was 26 children (22 boys and 4 girls) between the ages of 7 to 12 years old who were diagnosed by a physician or psychologist as having ADHD. The second group was comprised of 31 children (17 boys and 14 girls) who were 7 to 12 years old, except for one 15 year old. This second group was selected as a "normal" comparison group based on the parental report that these individuals were not identified to have any neurological, learning, emotional, or ADHD related problems. All subjects were volunteers.

The test instruments used included the IVA CPT, the Gordon CPT, the TOVA CPT, the CPRS-39 ADHD rating scale and the Children's Attention Scale (CAS). A standardized procedure was utilized in administering the CPTs, as described in their respective manuals. The CPTs were given in a counterbalanced order to control for fatigue effects. The rating scales were completed by one parent.

Diagnostic discriminative validity was evaluated by comparing the accuracy of the IVA CPT based classification using cut-off scores in its Interpretive Flowchart to diagnoses made by a physician or psychologist who had independently evaluated the child previously. IVA CPT's overall accuracy was found to be significant ($p<.0001$).

Table 1a. IVA CPT Classification of ADHD and non-ADHD groups

	Clinical Dx Of ADHD	Parent Rating Of Non-ADHD
IVA CPT Indicates ADHD	24	3
IVA CPT Indicates Non-ADHD	2	28
Total	26	31

Table 1b. Clinical Accuracy of the IVA CPT Test for ADHD

	Clinical Accuracy
Sensitivity	92%
Specificity	90%
Positive Predictive Power	89%
Negative Predictive Power	93%

The Sensitivity (proportion of ADHD children who are found positive on the measure) of IVA CPT was 92%. The Specificity (proportion of non-ADHD children who received a negative finding on the measure) of IVA CPT was 90%. The Positive Predictive Power (the proportion of test positives that have ADHD) of IVA CPT was 89% and the Negative Predictive Power (the proportion of test negatives that do not have ADHD) was 93%.

Concurrent validity was examined by comparing the children identified by IVA CPT likely to be ADHD in relationship to those ADHD children which were identified as positive by the other diagnostic instruments.

Table 2. Percent of Agreement of IVA CPT's Classification of ADHD with the Dx of ADHD made using other test instruments

	TOVA+	GORDON+	CPRS-39+	CAS+
IVA CPT+	90.0%	100.0%	91.7%	100.0%

The comparative accuracy of these various clinical instruments was assessed by examining their false negative rates. The criterion reference for this comparison of clinical accuracy was the pre-study clinical diagnosis.

Table 3. Percent of False Negatives by Test Instruments when compared to the Dx of ADHD made by Clinicians

	IVA CPT	TOVA	GORDON	CPRS-39	CAS
False Negative	7.7%	12.5%	36.0%	45.5%	59.1%

These results demonstrate that the IVA CPT has sufficient sensitivity (92%) and positive predictive power (89%) to be clinically useful in the diagnosis of ADHD in children on an individual basis. In the case of normal children populations, IVA CPT has an acceptable rate of false positives (less than 10%). IVA CPT had the lowest rate of false negatives among these test instruments, which previous research had shown was a major weakness of visual only CPTs. IVA CPT was also found to have excellent concurrent validity for both CPTs and parental ADHD rating scales. This validity research supports the conclusion that IVA CPT is an accurate psychological test which can provide important objective data as part of a comprehensive evaluation of ADHD with children.

The differential diagnostic accuracy of IVA CPT for other types of psychological problems has not been fully explored. However, clinicians have numerous psychological tests which can be used to rule out alternative diagnoses that may show as positive on the IVA CPT or to establish the co-morbidity of these diagnoses. In conclusion, this IVA CPT validity study provides evidence that supports the value of including auditory test stimuli in addition to the visual modality as an important element in the design of CPTs.

IVA CPT and ADHD Rating Scale Clinical Validity Study

The clinical usefulness and diagnostic validity of combining the IVA CPT and ADHD Rating scales was explored in a study of a clinical population of equal sex and matched ages over a wide age range. While the IVA CPT provides an objective measure of attention and response control in a quiet, structured environment, ADHD rating scale scores can help clinicians evaluate an individual's functioning in school, home and social/work environments. More distractions and off-task behavior can more easily occur in these settings. In addition, many individuals who are hyperactive will often show these types of ADHD symptoms when required to sit still over longer periods of time. Consequently, an evaluation of ADHD symptoms may prove clinically more accurate when the IVA CPT is used in conjunction with ADHD rating scales as part of a comprehensive evaluation for the typical types of clients seen in clinical practice.

In this study thirty clients previously diagnosed by Dr. Sandford as having a primary diagnosis of ADHD were selected. His diagnosis was based on a full evaluation including comprehensive psychological testing and ADHD rating scales. These clients had volunteered and given permission to have their clinical test data used in this study. These data records were de-identified. The patient data used in this study were selected from patients seen between the years of 2001 and 2011. Fifteen of these individuals were male, and fifteen were female. The ages ranged from 6 to 55, and the mean age was 17.1. Most of these individuals were diagnosed solely as having ADHD, but a few had a secondary diagnosis of Cognitive Disorder, Not Otherwise Specified. In these cases they were found to have significant cognitive processing problems in addition to their attentional deficits. Thirty individuals from the normative database, matched by age and sex, were randomly selected for comparison to this group. None of these individuals from the normative group were identified as having been diagnosed with ADHD, having ADHD-type symptoms, or having any other factors likely to impair their test functioning.

The ADHD classification was derived from the IVA-2 using the flowchart diagnostic algorithm and then compared to the clinical diagnosis. The clinical dagnosis was based on a full evaluation including a diagnostic intake evaluation and comprehensive psychological testing; including ADHD rating scales and a clinical interpretation of the IVA CPT. A comparison of the clinical diagnosis and the IVA CPT ADHD classification is presented below in Table 1a and 1b.

Table 1a. Comparison of Clinical Diagnosis and IVA CPT Classification

		Clinical Diagnosis		
		ADHD	**No ADHD**	**Total**
IVA CPT Classification	**ADHD**	24	5	29
	No ADHD	6	25	31
	Total	30	30	60

Table 1b. Clinical Accuracy of the IVA CPT Test for ADHD

	Diagnostic Accuracy
Sensitivity	80%
Specificity	83%
Positive Predictive Power	83%
Negative Predictive Power	81%

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.

The diagnostic accuracy of the IVA CPT test by itself showed that it was able to correctly identify about four out of five individuals who had been clinically diagnosed as having ADHD. The results in this second study also showed that the test misdiagnosed about one out of five individuals who did not have ADHD as having ADHD. Generally, rating scales are used by many clinicians to provide data in determining a diagnosis of

ADHD. Thus, it is useful to compare the accuracy of the IVA CPT test to that of rating scales reported in the research literature in order to clinically evaluate its comparative usefulness.

In evaluating individuals who sought treatment for clinical problems, Snyder, *et al.,* (2008) found that two different types of parent and teacher ADHD rating scales widely ranged in their diagnostic accuracy compared to clinicians' diagnoses which were based on a comprehensive, in-depth evaluation. The overall accuracy of ADHD rating scales for the clinical population in Snyder's study was ranged from 47% to 58%. This study's ADHD rating scale sensitivity in accurately diagnosing individuals with ADHD ranged from 38% to 78%. His results also found that the rating scales often mislabeled individuals as having ADHD who had either no diagnosis or another type of disorder. The low rates of specificity for this study ranged from 14-61%. A number of studies were also reviewed by Snyder that compared the accuracy of differentiating non-clinical populations from individuals diagnosed with ADHD, and the overall accuracy of these nine studies was higher, ranging from 55% to 79%.

The test results supported the clinical efficacy of the IVA CPT test by itself compared to ADHD rating scales. Based on the above review of the accuracy of ADHD rating scales, the IVA CPT test sensitivity of 80% is equivalent or better than the accuracy reported in most of the ADHD rating scale studies. The specificity of the IVA CPT (83%) is actually better than the highest accuracy rate for identifying individuals who do not have ADHD when compared to the ADHD rating scale research discussed above. In general, it has been noted that rating scales often have low rates of specificity in that many individuals without ADHD are misclassified as having ADHD.

The rating scale classification was also compared to the clinical diagnosis. In this case since rating scale data was not available for the matched normative sample group, only sensitivity and negative predictive power can be reported. The sensitivity of the rating scales alone was equal to the sensitivity for the IVA CPT by itself, and the negative predictive power of the rating scales was slightly higher than that of the IVA CPT. (See Tables 2a and 2b presented below.)

Table 2a. Comparison of Clinical Diagnosis and ADHD Rating Scales

		Clinical Diagnosis		
		ADHD	**No ADHD**	**Total**
ADHD Rating Scale Classification	**ADHD**	23	0	23
	No ADHD	7	30	37
	Total	30	30	60

Table 2b. Clinical Accuracy of ADHD Rating Scales

	Classification Accuracy
Sensitivity	77%
Negative Predictive Power	81%

In this validity study, the question was also addressed as to whether the IVA CPT test results in combination with the rating scales may further help clinicians in accurately diagnosing ADHD. Consequently, the rating scale diagnosis for each client was determined by using the symptom cut-off guidelines for hyperactive/impulsive and inattentive symptoms appropriate to the client's age (American Psychiatric Association, 2013). Parent, teacher, and self-rating scales were combined. If any one of the available scales had ADHD symptoms above the cut-off for either hyperactive/impulsive or inattention symptoms, that positive symptom rating was used in formulating a diagnosis. In other words, if the parent rating scale identified six hyperactive/impulsive symptoms and the teacher identified six inattentive symptoms then the rating scales were interpreted as supporting the diagnosis of ADHD, Combined presentation.

The IVA CPT and rating scale data were combined in diagnosing ADHD. If either of these two methods supported a diagnosis of ADHD, then that diagnosis was assigned. There was no rating scale data for the normative population, so for the normative population, only the IVA CPT results were used in determining a diagnosis of ADHD. The results of the IVA CPT and rating scale diagnoses in comparison to the clinical diagnosis are presented below in Tables 3a and 3b.

Table 3a. ADHD Clinical Dx compared with the IVA CPT and ADHD Rating Scale Classification

		Clinical Diagnosis		
		ADHD	**No ADHD**	**Total**
IVA CPT & Rating Scale Classification	**ADHD**	27	5	32
	No ADHD	3	25	28
	Total	30	30	60

Table 3b. Clinical Accuracy of the IVA CPT Combined with ADHD Rating Scale

	Clinical Accuracy
Sensitivity	90%
Specificity	83%
Positive Predictive Power	84%
Negative Predictive Power	89%

The combination of the IVA CPT test results and the rating scales increased the overall sensitivity by 10%. The specificity remained the same. The positive predictive power increased by 1%. The negative predictive power improved by 8%.Clinicians using the IVA CPT in combination with ADHD rating scales would be able to classify individuals with ADHD and without ADHD with an overall accuracy of 90%.

The value of including the rating scales in making a diagnosis of ADHD is likely to be due to the fact that they provide a measure of the occurrence of gross-motor hyperactivity that is not specifically identified by the IVA CPT test. In addition, the rating scales provide data relevant to the individual's functioning in both the home and school

environments in respect to ADHD symptoms which may not manifest under the more controlled test conditions required for administering the IVA CPT. In contrast, the test results provide the clinician with the opportunity to objectively measure clients' mental processing speed and its variability, attentional functioning, and impulsive responses that may be aspects of ADHD that are difficult for raters to accurately identify in the work and school environments. Separately, the rating scale and the IVA CPT are equivalent in respect to their accuracy in classifying ADHD, but in combination, they were found to be more accurate in this clinical population study. These findings support combining the IVA CPT test results with rating scales in helping clinicians to make more accurate diagnoses of ADHD when working with clinical populations of all ages.

This validity study differs from the first validity study discussed above in a number of ways. This first study examined the validity of the IVA CPT by itself in classifying children suspected of having ADHD. The subjects in the first study were mostly boys (85%) and the age range was 7 to 12 years old. In the second validity study subjects were selected so that number for each sex was equal and the age range was much wider (ages 6 to 55). The Sensitivity in the first study was 92% in comparison to the IVA CPT's Sensitivity with a mixed age clinical group of 80%. It is likely that this difference is due in part to what appears to be the greater sensitivity of the IVA CPT in identifying ADHD symptoms of young boys. In addition, the second study included subjects who had presented for general clinical treatment and not just for ADHD, but with ADHD-type symptoms pertaining to other diagnoses. Thus, the validity for using both the IVA CPT and ADHD rating scales with clinical populations of all ages is evident in this study by its clinical sensitivity of 90% in identifying ADHD and at the same time the ability to classify 89% of individuals who did not have ADHD.

IVA CPT Test-Retest Reliability Study

Joseph A. Sandford, Ph.D. completed a reliability study of IVA+Plus in conjunction with Philip Seckler, M.S., William Burns, Ph.D., and Doil Montgomery, Ph.D. from NOVA Southeastern University. This study was included as part of Dr. Seckler's dissertation.

Reliability refers to the consistency of test score performance for repeated testing by an individual under similar conditions. A test-retest reliability study of IVA CPT was completed by BrainTrain in conjunction with NOVA University. This type of reliability test provided an index about the stability of IVA CPT test scores over time (Anastasi, 1988) that allows therapists to be confident that the changes observed in scores reflect differences in a person's performance and are not solely due to random errors. If the IVA CPT test results are found to be consistent over time, then they can be practically used in clinical diagnostic decision making and in the evaluation of medication and/or treatment effects. The purpose of this study was to determine the test-retest reliability of the 22 raw scale scores and the derived six composite quotient scales. The composite quotient scores are based on the raw scores of selected relevant scales and statistically derived.

A total of 70 individuals without identified problems of neurological, current psychological, learning, attention or self-control problems were given the IVA CPT on two separate occasions. The age range of volunteers was 5 to 70 years old and the mean age was 21.8 years. Sixty percent of these individuals were females and 40% were males. All subjects were volunteers.

A standardized procedure was utilized in administering the IVA CPT, as described in its manual. Testing was 1 to 4 weeks apart. Most of the IVA CPT test instructions are presented visually on the monitor screen and then spoken in a clear, digitized female voice. A warm-up session for both auditory and visual targets was first given. All participants were given this opportunity to learn how to click the mouse correctly. Next, each participant completed a 32 item practice session to learn with the examiner's help (if necessary) how to respond correctly when both targets (**1**'s) and foils (**2**'s) were presented in a mixed up order. If during the practice session the test taker had problems doing the test task correctly, then the examiner could temporarily stop the practice training and explain the four simple test rules which were:

1. Click when you see a "**one"**
2. Click when you hear a "**one"**
3. Don't click when you see a "**two"**
4. Don't click when you hear a "**two**"

Once the main part of the IVA CPT test began, no further instructions could be given, except to redirect the test taker if he or she removed his/her finger from the correct mouse button. At the end of the test a second simple reaction time test just like the warm-up section was given again. Scores were then automatically saved by the computer for later analysis.

Six tables are presented below which summarize the data of this reliability study. Statistical analysis was completed to determine whether test scores were similar for each individual for both testing sessions and whether any learning or improvement in test scores occurred.

Table 1.

The Means and SD of IVA CPT Composite Quotient Scores for Test 1 (T1) and Test 2 (T2)

Scale	Mean T1	Std Dev	Mean T2	Std Dev
FRCQ	101.12	13.53	103.41	13.50
ARCQ	101.94	13.70	103.32	14.25
VRCQ	100.17	13.58	102.88	14.77
FAQ	104.43	10.85	106.72	12.97
AAQ	104.54	11.07	105.64	14.04
VAQ	103.69	12.10	106.83	11.88

Table 2.

The Means and SD of the 22 IVA Raw Scale Scores for Test 1 (T1) and Test 2 (T2)

Scale	Mean T1	Std Dev	Mean T2	Std Dev
HYP	6.09	9.89	5.61	10.72
CMPA	99.50	0.99	99.13	1.83
CMPV	99.17	2.30	98.89	2.22
CONA	72.17	7.41	71.27	7.10
CONV	69.73	5.21	71.60	5.24
FOCA	71.23	7.02	71.34	8.09
FOCV	71.29	6.12	71.87	7.14
MNA	479.16	106.02	460.59	113.40
MNV	343.31	88.89	332.14	96.07
PRA%	94.00	6.22	95.64	5.44
PRV%	94.10	7.48	94.20	7.39
RFRA	91.10	10.07	92.01	11.56
RFRV	88.00	9.74	86.06	10.86
RVAC	71.51	7.75	72.09	9.44
RWCA	100.67	40.92	103.93	39.07
RWCV	92.53	21.13	99.09	31.02
SMA	177.96	85.41	164.23	83.46
SMV	143.99	48.81	140.34	46.55
STMA	94.04	9.30	95.19	9.47
STMV	96.30	9.77	96.81	8.57
VIA%	98.87	2.55	98.84	2.33
VIV%	98.16	4.13	98.36	4.32

Table 3.

The Pearson *r* for the Test-Retest of the IVA CPT Composite Quotient Scores

Scale	Correlation (r)	Sig. 2 tail
FRCQ	0.41	p<.01
ARCQ	0.39	p<.01
VRCQ	0.37	p<.01
FAQ	0.74	p<.01
AAQ	0.66	p<.01
VAQ	0.75	p<.01

Table 4.

The Pearson *r* for the Test-Retest of the 22 IVA CPT Scale Raw Scores in Rank Order of Lowest to Highest Correlations

Scale	Correlation (r)	Sig. 2 tail
RWCA	0.02	n.s.
STMV	0.18	n.s.
STMA	0.26	p<.05
VIA%	0.32	p<.01
RWCV	0.34	p<.01
RFRA	0.46	p<.01
RFRV	0.47	p<.01
CONV	0.52	p<.01
SMA	0.58	p<.01
PRV%	0.61	p<.01
PRA%	0.64	p<.01
FOCV	0.65	p<.01
FOCA	0.68	p<.01
CONA	0.68	p<.01
RVAC	0.69	p<.01
CMPA	0.70	p<.01
SMV	0.70	p<.01
VIV%	0.71	p<.01
HYP	0.80	p<.01
CMPV	0.80	p<.01
MNA	0.87	p<.01
MNV	0.88	p<.01

Table 5.

The T-Test Probability and Percent Change for the Test-Retest of the IVA CPT Composite Quotient Scores

Scale	Diff T2 – T1	Change %	Prob.
FRCQ	2.29	2.26%	0.195
ARCQ	1.38	1.35%	0.456
VRCQ	2.71	2.71%	0.159
FAQ	2.29	2.19%	0.033
AAQ	1.10	1.05%	0.391
VAQ	3.14	3.03%	0.003

Table 6.

The T-Test Probability and Percent Change for the Test-Retest of the 22 IVA CPT Scale Raw Scores in Rank Order of Lowest to Highest Probabilities

Scale	Diff T2 – T1	Change %	Prob.
CONV	1.87	2.68%	0.003
PRA%	1.64	1.75%	0.008
MNA	-18.57	-3.88%	0.008
CMPA	-0.37	-0.37%	0.024
CONA	-1.50	-2.06%	0.034
MNV	-11.17	-3.25%	0.043
RWCV	6.56	7.09%	0.083
CMPV	-0.29	-0.29%	0.098
RFRV	-1.94	-2.21%	0.131
SMA	-13.73	-7.71%	0.143
FOCV	0.59	0.82%	0.388
STMA	1.14	1.22%	0.406
SMV	-3.64	-2.53%	0.410
RVAC	0.57	0.80%	0.492
RFRA	0.91	1.00%	0.500
HYP	-0.47	-7.75%	0.550
VIV%	0.20	0.20%	0.608
RWCA	3.26	3.24%	0.636
STMV	0.51	0.53%	0.716
FOCA	0.11	0.16%	0.877
PRV%	0.10	0.11%	0.898
VIA%	-0.03	-0.03%	0.933

The first way to assess test reliability is to determine if the test scores significantly change when the same person is re-tested. The IVA CPT Reliability study shows very small practice effects for the subjects who completed the IVA CPT on two separate occasions. Only the Full Scale and Visual Attention quotient scores significantly changed ($p<.03$ and $p<.003$, respectively). The other four global scales showed no significant practice effect change. It needs to be noted that the Full Scale Attention (FAQ) quotient score improved by a mean of 2 points and the improvement for the Visual Attention (VAQ) quotient was 3 points on average. This increase of between two to three quotient points is the same practice effect reported in various studies for the WISC-III Verbal IQ test scores. The WISC-III Performance IQ test scores had a much greater practice effect of 11-13 points. Thus, the IVA CPT quotient scores showed very little change on the second test administration and the change observed was less than what has been observed in re-testing of IQ.

A second way to analyze a test's reliability is to examine the scale score correlations between the two test administrations. All of the correlations between the two IVA CPT test administrations were significant for the Global scales and for all except one of the key primary scales. These correlations were rated as Strong for the Global Attention scales FAQ, AAQ and VAQ (.66 to .75) and mild to moderate for the Response control scales (.37 to .41). Almost all of the main key primary scales had strong correlations (.52 to .88), but the two Stamina scales both had weak correlations (.18 to .26). On the second test administration the subjects probably lost interest in maintaining their effort and were slower in their reaction time as the test proceeded. It was the Stamina scale that loads on the Global Response Control scales and accounts for being lower. Thus, the scale quotient score changes and the overall correlation test re-test findings support that the IVA CPT has good test reliability.

All IVA CPT composite quotient scores showed significant correlations for test-retest scores, demonstrating the stability of this test over time. The correlations ranged from .37 to .75. The response control quotients had correlations which showed a moderately strong positive relationship and the attention quotient score correlations demonstrated very strong positive relationships. Given that the scores for many of the IVA CPT primary scales for this "normal" population showed very low error rates (1% to 6%), the relatively high correlations obtained further reflect the stability of the IVA CPT test. The analysis of the 22 IVA CPT scale raw scores found 20 scales had significant positive relationships and 18 out of these 20 correlations showed a moderately strong to very strong relationship (.46 to .88). Thus, the IVA CPT was found to be a significantly stable measure of performance in many ways both globally and in terms of specific scales.

The overall changes in quotient scores were very small ranging from 1.05% to 3.03%. Nonetheless, a statistically significant (p<.01) improvement of 3.03% in the Visual Attention Quotient (VAQ) was found. The VAQ increased from 103.7 to 106.8. Examination of the sub-scales which load on the VAQ showed that a small improvement in visual reaction time (3.25%) occurred which led to the increase. No other quotient scores showed significant learning effects. Very small, but statistically significant improvements (p<.01) were found for 3 of the 22 raw score scales. These increases showed very small changes in the Mean Auditory (MNA) reaction time (-3.88%), Prudence Auditory (PRA%) response inhibition (1.75%) and in the Consistency of Visual (CONV) reaction times (2.68%). Given that these changes in scores were very small and relatively few, it is concluded that almost all individuals who are motivated to perform well and whose mental and psychological states have not changed as a result of factors recognized to affect concentration, attention and response inhibition (e.g., fatigue, illness, stress, etc.) will not show any substantial practice or learning effects on repeated administrations of IVA CPT. Thus, changes which occur in IVA CPT scores over time can reliably be interpreted to reflect possible medication, treatment or environmental effects.

Further in-depth analysis found evidence that individual scales related to stamina (STMV and STMA) and persistence (RWCV and RWCA) were the least reliable. Stamina reflects a change in discriminatory reaction time speed over the test and persistence involved a comparison of simple reaction time speed measured before and after the main IVA CPT test. These scales can be interpreted as reflecting an individual test taker's state. When a person is tired or poorly motivated these scale scores may reflect this state and vice-versa. From the opposite perspective (i.e., traits), the mean speed of discriminatory reaction time for both visual and auditory correct responses (MNV and MNA) were the two most stable measures identified. Discriminatory reaction time may reflect a stable characteristic for most individuals without attention problems, possibly related to innate mental processing speed.

Thus, IVA CPT offers the possibility to assess both state and trait characteristics based on the comprehensiveness of its scale analysis. In other words, this feature of IVA CPT makes it a sensitive test of both motivation and ability.

Normative Database Study

Computerized Continuous Performance Tests (CPTs) in the past have typically presented visual stimuli, making the assumption that tests in only one sensory modality can adequately detect the problems of general attentional and inhibitory control processes. The implied assumption of visual only CPTs is that there are no significant differences between visual and auditory modalities. Recent research by Taylor (1994) found that, in a normal college population, significantly more errors of commission occurred in a non-computerized auditory CPT as compared to a visual CPT task. The recent development of multi-media computers now makes it possible to comprehensively assess auditory as well as visual processes with a computerized CPT, in order to replicate and extend this finding. Thus, this study's major purpose was to clarify the possible need and value for including both visual and auditory sensory modalities in CPT evaluations.

Similarities and differences of performance on an auditory and visual CPT for different age and sex groups were also explored in this normative study. Sex and age group differences would support the need for normative groups to be provided by appropriate age and sex groups for accurate clinical interpretation. Consequently, a secondary purpose of this research was to evaluate the normative data collected with the Integrated Visual and Auditory (IVA) CPT to determine what possible age and sex differences exist.

The normative database used in this study consists of 487 individuals (210 males and 277 females) ranging in age from five to 90 years. The current normative database was updated to 1700 individuals in December, 1999, but was not used in this study. These volunteers were without identified neurological, current psychological, learning, and attentional or self-control problems.

A standardized procedure was utilized in administering the IVA CPT, as described in its manual. Once the main part of the IVA CPT test began, no further instructions could be given, except to redirect the test taker if he or she removed his/her finger from the correct mouse button. At the end of the test, a second simple reaction time test just like the warm-up section was given again. Scores were then automatically saved by the computer for later analysis.

Auditory and visual performance was found to be highly correlated ($p < .001$) for each of the six primary scales. Thus, subjects who performed well in one sensory modality tended to perform well in the other modality. The correlations for the six primary scales ranged from .53 to .86; demonstrating moderate to very strong relationships.

Errors of commission and omission were not found to be randomly distributed. Errors of commission for both auditory and visual modalities occurred significantly more

frequently (p <.001) when the targets ("**1**"s) were common and the subject had to inhibit his responses to the infrequent, interspersed non-targets ("**2**"s). Likewise, errors of omission for both visual and auditory modalities were observed to be significantly (p < .001) more when the non-targets ("**2**"s) were common and the subject had to respond to the infrequent, interspersed targets ("**1**"s).

Significant gender differences were found on two scales. Mean male reaction times for correct responses were 23 milliseconds faster than females (p < .02). Females showed significantly fewer errors of commission (3% less) than males (p < .001).

Mean reaction time for correct responses (i.e., the Speed scale) by age generally followed a U-shaped curve. A rapid improvement (i.e., reduction in reaction time) was observed between the ages of 5 and 7, indicating that the test task is quite demanding for younger children. The ability to switch attention between two sensory modalities may be a developmental milestone. A more gradual reduction in reaction time followed for children eight to twelve years old, reaching optimum performance in the mid-teen to young adult years (perhaps not coincidentally the age when driver permits are issued). Reaction time was fairly stable through middle age and then began to slow down slightly from the age of 45 on.

Four of the primary scales were discovered to show differences between the two sensory modalities. These test findings are listed below:

The scores on the Prudence scale showed that auditory errors of commission (Mean = 7.9%) were found to be significantly (p < .01) more frequent than corresponding visual commission errors (Mean = 6.9%).

Scores on the Vigilance scale demonstrated that visual errors of omission (Mean = 5.8%) had a significantly higher rate of occurrence (p < .01) than auditory errors of omission (Mean = 2.9%).

Stamina scale scores indicated that the auditory processing showed more mental fatigue, as the speed of reaction time significantly (p < .001) decreased more for auditory trials (Mean = 6.2%) than for visual trials (Mean = 2.8%) over the course of the test.

Based on the Consistency scale, the visual reaction times (Mean Consistency = 69.3%) were found to be significantly (p < .001) more variable than auditory reaction times (Mean Consistency = 70.9%).

The high significant correlations found between all auditory and visual scales suggested that there is an underlying attentional ability which generalizes to both modalities.

However, given that these correlations were not perfect, some individuals may be relatively dominant in one or the other modality. In other words, some people are likely to have strengths or weaknesses in either visual or auditory modalities.

The results of this study confirmed the theoretical basis of the IVA CPT test design. During the trials when the "**1**"s were more frequent, commission errors occurred much more frequently for both sensory modalities. Likewise, when "**2**"s were more frequent, errors of omission were more prevalent.

Several significant differences for auditory and visual attention processing by age and sex were found using the IVA CPT. While males had faster reaction times than females, females made fewer impulsive errors. These two factors can make mutually exclusive demands; in other words, it may be difficult to be both quick and accurate. This finding may be related to nature and nurture issues; requiring further research. While this study does not clarify the causes of this difference in the sexes, it does suggest that males in this CPT were more predisposed than females to react more quickly and, thus, made more impulsivity errors.

The reaction time findings related to age suggested that there may be attentional developmental milestones in terms of speed of mental processing which can be measured quantitatively. Based on this reasoning, it may be useful to consider various factors of attention (i.e., being vigilant, focused and quick to respond) and response control (i.e., inhibiting impulsive erroneous responses, being consistent and maintaining response speed) in helping to understand and measure the maturity of mental development. Research examining the correlation of IVA CPT scales, general intelligence, academic achievement and behavioral problems may also prove valuable in increasing our understanding of the human mind's abilities and disabilities.

In general, these results found a normal population to be more aurally impulsive and visually inattentive within the IVA CPT paradigm. This finding suggests that people may exhibit more of a startle response to auditory stimuli, which possibly then evokes a reflex reaction, even when it is inappropriate. It also appears that inattention errors to visual stimuli may be higher than auditory stimuli due to the necessity to maintain almost constant visual focus on the computer screen in order to avoid missing a target and because visual stimuli are presented for about one third of the time of auditory stimuli. These characteristics of the visual stimuli in IVA CPT may also account for the greater variability in visual reaction time found in comparison to auditory responses. Auditory processing reaction time was observed to fatigue more than visual responses during this short test.

In conclusion, the significant findings of this study supported the need for CPT tests to include both sensory modalities in order to comprehensively evaluate Vigilance, Prudence, Consistency and Stamina. The age and sex differences also found in this study support that the clinical interpretation of IVA CPT test scores will be most accurate when compared to the appropriate sex and age normative group.