TEST-RETEST RELIABILITY AND VALIDITY OF THE COMPUTERIZED VERSION OF THE CATEGORY TEST-YOUNG CHILDREN'S VERSION

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTORATE OF PHILOSOPHY

IN THE GRADUATE SCHOOL OF THE

TEXAS WOMAN'S UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

BY

DEBRA J. KELLEY-GOMEZ, M.S.

DENTON, TEXAS

AUGUST, 1998

TEXAS WOMAN'S UNIVERSITY
DENTON, TEXAS

June 24, 1998

To the Associate Vice President for Research and Dean of the Graduate School:

I am submitting a dissertation written by Debra J. Kelley-Gomez entitled "Test-Retest Reliability and Validity of the Computerized Version of the Category Test-Young Children's Version." I have examined this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctorate of Philosophy with a major in School Psychology.

Daniel C. Miller, Ph.D., Major Professor

We have read this dissertation and recommend its acceptance:

Department Chair

Accepted

Associate Vice President for Research
and Dean of the Graduate School

This is dedicated to my grandparents....

Thanks

Grandma and Grandpa you gave me unconditional love

Jose you gave me a best friend and love

Kids you gave me my soul

Mom and Dad you gave me my wings

Sisters you gave me perspective

Jayne you gave me a fresh look at life

Lupe and Joe you gave me family

Erik you gave me my eternal subject

Alisa you gave me spirit and my kids a Na Na

Carol you gave me unwavering support

Tom you gave of yourself to my family

Kim you gave me friendship

Dr. Day and Dr. Aspy you showed me more than you will ever know

Dr. Miller, Dr. Vitro, and Dr. Graham, Thank you for your encouragement and support.

"Are we there yet?"

Test-Retest Reliability And Validity Of The
Computerized Version Of The Category Test-Young Children's Version

Debra J. Kelley-Gomez

August, 1998

The primary focus of the research involved the reliability and validity of the Computerized Version of the Category Test -Young Children's Version (CVCT-YC). There were 56 elementary school children, ages 5-8 (24 females 42.9% and 32 males 57.1%) who participated in this study. The subjects were drawn from several YMCA after-school programs in northern Texas and the mean age was 6.78 years.

Statistical results indicated that the CVCT-YC is indeed statistically stable over time, specifically over three-week periods. The CVCT-YC was also found to correlate significantly with a psychometrically established measure of fluid reasoning: the Concept Formation Test, which originates as a subtest from the Woodcock-Johnson Tests of Cognitive Ability. This research also examined for significant relationships between particular populations and the results of either the CVCT-YC or Concept Formation assessments. This research examined for the possible effects of age, gender, hand dominance, ethnicity, income, mother's highest level of education, father's highest level of education, primary caregiver's marital status, and with whom the subject lives. None of the populations were found to exhibit a significant relationship with assessment scores.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

CHAPTER I

Literature Review

In the past quarter century neuropsychology has developed into a field of widespread recognition and acceptance as a formal applied professional specialty area. Its origins as a scientific field of inquiry date from the late 19th century. Until recently, neuropsychology was primarily identified with the diagnostic testing of adults with verified brain injury. With the increasing recognition of neuropsychological substrates of adaptive behavior problems and learning in adults with brain injury, there developed an advancing interest in the possibility that central processing dysfunctions may be related to a wide variety of children's learning problems (Reynolds & Fletcher-Janzen, 1997).

Progressively, through programs on the identification, description, and treatment of childhood learning problems, neuropsychology became involved with the assessment of exceptional children. Since the body of research relating known brain damage to specific behavior and learning problems had for the most part involved adults, the question evolved concerning the extent to which that research could be applied to children. The professional issue that stemmed from this scientific question was, namely, which tests or diagnostic approaches were appropriate for use with children (Hartlage & Long, 1997).

In attempts to address the scientific and professional issues raised by the demand for neuropsychological services for children, two diverse approaches of neuropsychological services to children emerged. The first approach involved modified versions of traditional neuropsychological batteries, such as the Halstead-Reitan Neuropsychological Battery (Reitan, 1955; Reitan & Davidson, 1974; Selz, 1981) and the Luria-Nebraska Neuropsychological Battery (Golden, Hammeke, & Purisch, 1980; Plaisted, Gustavson, Wilkening, & Golden, 1983), which standardized the adult battery items on a child sample. The standardization took the form of deleting from the adult battery those items that were too difficult for children (Hartlage & Long, 1997).

The second approach placed emphasis on the interpretation of standard psychometric tests from a neuropsychological perspective using appropriate tests for children of given ages ranging from preschool to adolescence (Hartlage & Long, 1997; Hartlage, 1981, 1984; Hartlage & Telzrow, 1983; Telzrow & Hartlage, 1984). This approach implemented standardized behavioral tests and interpreted them according to the individual's strengths and weaknesses. Such an approach was popular with psychologists working in school settings, because they were likely to encounter trouble justifying the addition of a neuropsychological battery to the standard array of psychometric instruments already required by most schools for psychoeducational assessment.

Recently, some practitioners have begun to modify subparts of existing neuropsychological batteries, such as the "revised" Children's and Young Adolescent's Version of the Category Test (Boll, 1993) to assess for any indication of brain damage. This paper will focus on the Computerized Version of the Category Test-Young Children's Version (CVCT-YC) and its development, application, validity, and reliability.

As there is very little published data on the CVCT-YC, much of the review of literature will cover research published concerning related versions of the original Category Test, such as the Category Test-208 (CAT-208) or the Children's Category Test (CCT). The Children's Category Test is a modified version of the CAT-208, and the CVCT-YC is a modified version of the CCT. An example of this modification process is given later in this paper. For this reason, each version of the Category Test must exhibit a strong foundation of psychometric properties in order to be built upon.

For purposes of introduction, the following areas will be examined: (a) the Category Tests, (b) the Children's Category Test, (c) reliability, (d) validity, (e) automation of the Category Tests, (f) The Computerized Version of the Category Test-Young Children's Version, and (g) the purpose of the current research.

The Category Tests

The Original Category Test. During a 1940 study, W.T. Halstead attempted to better correlate aspects of psychological behavior with specific cerebral anatomy. He presented a set of 62 actual objects to individuals with definite, localized brain lesions (most had excised brain sections). The subjects were to: 1) group the objects according to a principle or rule, 2) respond to questions whether one object or another could be removed or added without violating the principle or rule, 3) respond to questions about groupings made by the examiner using the same objects, and 4) freely recall which objects were used in testing.

Halstead found that those with frontal lesions, left or right, showed characteristically different grouping behavior than did his normal controls or other brain-damaged subjects.

The frontal-lobe-damaged subjects showed fewer objects used in spontaneous groupings, fewer objects in recall, and fewer numbers of groupings, despite the fact that they took more time to perform the tasks. On the original Category Test, no feedback was given to the subjects concerning correct or incorrect responses, and the scoring and presentation were never standardized to a useful level. The original Category Test was designed as a measure of an adult's ability to learn from experience while adaptively integrating new information, and was intended to be largely independent of culture, language, and education (Boll, 1993). According to Halstead, the inability to perform these tasks was a strong indicator of brain damage (Halstead, 1940).

In 1943, Halstead and Sattlage created a new version of the test. It retained the principles of his original version but attempted to be less subjective and more quantifiable. Halstead presented geometric figures to the subject by means of a revolving drum. Attached to the drum was a piece of white cloth, on which figures were drawn. A subject could only view one figure set (four drawings) at a time. Subjects were instructed to flip a switch corresponding to a number from 1 to 4 whatever they felt the picture suggested. Correct and incorrect feedback for this test was introduced. If the subject chose correctly, the drum rotated to the new figure set. If an incorrect response was made, nothing happened. If subjects were having trouble, they were allowed to be coaxed or even given the principle. The test comprised nine subtests. Each subtest was created according to a principle: either number, oddity, part-whole, or recognition. The same principles remain in use today.

In 1947, Halstead selected a series of seven tests based on their perceived power to differentiate patients with frontal neuropathology from either non-frontally involved patients or normal controls. The Category Test comprised 336 visually presented stimulus figures in which the subject's task was to abstract the underlying organizing principle. The initial 16 items enabled the subject to develop simple associations between Roman numerals and a number corresponding to one of four reaction keys. The individual learned in the sample test that positive reinforcement followed correct responses, and negative reinforcement followed incorrect responses. The feedback principle involved subjects either hearing a chime (correct response) or a buzzer (incorrect response).

Through factor analysis, Halstead found significant loading on what he called factors of central integration and abstraction ability. Central integration was considered the psychological function that compares new information with familiar ground information, and either accepts (incorporates) or rejects it. Abstracting ability was considered to be the function of grouping according to criterion. Thus, failures on the Category Test resulted from a deficit in central integration processing rather than from an inability to group to criterion.

CAT-208. Reitan and Davidson (1974) adapted the Category Test developed by Halstead (1947) for inclusion in their neuropsychological test battery. The Reitan and Davidson (1974) version (CAT-208) consisted of seven sets of items with a total of 208 items. Organization of each set of items was on the basis of an underlying principle such as number of objects, odd stimulus, ordinal position, or color (Spreen & Strauss, 1991). The basic purpose of the Category Test was to assess an individual's capacity to alter performance based on positive and negative reinforcement. The essential nature of the test was an experiment in abstraction or concept formation (Reitan & Wolfson, 1985; Spreen & Strauss, 1991).

Halstead (1947) considered the original Category Test to be a measure of abstraction. Reitan and Davidson (1974) described the CAT-208 test as evaluating "concept formation and the ability to apply organizing principles in performance of complex procedures" (p.58). The Category Test required an individual to assess and understand the nature of a problem situation. Then, the individual had to postulate solutions based on past trial and error. The test did not require verbal abilities; instead it demanded flexibility in abstraction and concept formation. Laatsch (1983) noted that the CAT-208 required metaknowledge, or "knowledge about knowledge." Bertram, Abeles, and Snyder (1990) reported that effective performance on the CAT-208 required skill in testing discrepancies among stimuli. Spreen and Strauss (1991) suggested that the CAT-208 required a special competence in abstraction abilities and that the CAT-208 demanded solution to a variety of problem situations. Today, the Category Test (CAT-208) is well-established as one of the most sensitive indicators of brain damage (Boll, 1978; Boll, 1993; Boll & Reitan, 1972; Doehring & Reitan, 1962; Golden, 1978; Halstead, 1947; Klove, 1974; Matthews, 1974; Mercer, 1993; Shore, Shore, & Pihl, 1970).

However, despite its unique clinical usefulness, the Category Test (CAT-208) was excluded from many assessments because of the practical limitations associated with its lengthy administration time, expensive equipment, and limited transportability across clinical settings (Boyle, 1986; Calsyn, O'Leary, & Chaney, 1980; Gregory, Paul, &

Morrison, 1979; Sherrill, 1987; Summers & Boyle, 1987; Taylor, Goldman, Leavitt, & Kleinman, 1984). For instance, administration time could exceed two hours, especially when evaluating a neurologically impaired individual. The test items were presented as a series of slides projected onto a glass screen. The examinee responded by choosing one of four levers, each of which was hooked to a bell and buzzer to signal positive or negative feedback.

Numerous researchers have investigated the feasibility of reducing both the administration time and overall complexity associated with the Category Test (Barker, 1977; Boyle, 1975, 1986; Calsyn et al., 1980; Golden, Kuperman, MacInnes, & Moses, 1981; Gregory, et al., 1979; Kimura, 1981; MacInnes, Forch, & Golden, 1981; Taylor et al., 1984). Kilpatrick (1970) first examined the feasibility of a short form of the CAT-208. In a sample of 41 patients, reported correlation coefficients for the total error score to the error score for odd items was .90, .99 for the total to even items, and .97 for the odd items to even items. This suggested that one could reduce the CAT-208 to 104 items and have a reliability coefficient of .94. Many of these studies were criticized for not including analyses of correlations among other instruments typically contained in a neuropsychological battery, leaving little evidence to support the test equivalency of these abbreviated forms or their ability to discriminate brain damaged from non-brain-damaged individuals (Boll, 1993).

Different administration forms of the category test. Today the Category Test exists in several different versions, including the standard projection version (CAT-208) (Reitan and Wolfson, 1985), the booklet version (BCT) (DeFilippis & McCampbell, 1979), and a computerized version (CVCT-A) (Miller, 1989). DeFilippis and McCampbell (1979) developed a booklet form of the CAT-208. The method of administration was the major change between the CAT-208 and the BCT. Other minor changes included a score sheet depicting the test items and instructions for administration.

DeFilippis and McCampbell (1979) examined the relationship between the CAT-208 and BCT using a sample of 30 normal college students; the reported overall correlation coefficient was .91. Oddly enough, those who received the BCT first had a reported correlation coefficient of .94. Those who received the CAT-208 first had a reported correlation coefficient of .89. A significant practice effect occurred between the initial and second administration of the Category Test. However, order of the test version administered was not found to be significant. DeFilippis, McCampbell, and Rogers (1979) also examined the relationship between the BCT and the CAT-208 in a sample of 30 chronic alcoholics. The reported overall correlation coefficient was .80. In the sample of alcoholics, for those who received the BCT first, the correlation coefficient was .75. For those who received the CAT-208 first, the correlation coefficient was .82. Not only was there significant evidence of practice effect, the order of the test version administered was significant. Alcoholics who received the BCT initially had lower mean scores than did those who received the CAT-208 initially.

MacInnes, et al. (1981) cross-validated the BCT. Male inpatients were divided into three diagnostic categories: organic brain syndrome ($\underline{n}$=16), schizophrenia without evidence of organicity ($\underline{n}$=13), and personality disorder ($\underline{n}$=9). The correlation coefficients between the CAT-208 and the BCT were: .78 for the organic brain syndrome group, .52

for the schizophrenia group, and .94 for the personality disorder group. The correlation with the CAT-208 of .74 suggested that the BCT could be an adequate substitute for the CAT-208 without significant loss of reliability or validity. DeFilippis and McCampbell (1979) suggest that the reason the BCT did not reliably measure the performance of schizophrenics is because test reliability, in general, is lower in patient populations.

Beaumont (1975) examined the validity of administering the CAT-208 via online computer. When a sample of 10 brain-damaged patients were tested, and 10 psychiatric patients, results indicated that the two groups did not differ in performance on the CAT-208. The online CAT-208 misclassified 30% of the brain-damaged group and 70% of the psychiatric group. With a cutoff of 50 errors, this made for a total misclassification rate of 50%. However, these findings were consistent with Goldstein and Shelly (1972) who used the traditional administration method of the CAT-208.

In 1989, Miller developed a computerized version of the Category Test for the Macintosh computer. Rockers (1996) examined the test-retest correlation of the Computerized Version of the Category Test by splitting his subjects into three groups and administering each a different version of the Category Test. Approximately one-third took the standard version ($\underline{n}$=11), one-third took the booklet version ($\underline{n}$=9), and the remaining third took the computerized version ($\underline{n}$=10). For the retest administration (approximately one to two months after the first), all subjects were given only the Computerized Version of the Category Test. When correlated by test-retest groups (CAT-CVCT-A, BCT-CVCT-A, CVCT-A-CVCT-A), the CAT-standard version-correlated the lowest and the BCT the highest. The CVCT-A ranked moderately between the CAT and BCT. Rockers

(1996) noted that a possible reason for the less-than-expected correlation of the CVCT-A could have been due to problems involving the subject clicking the mouse twice in order to respond to a computer question instead of once, thus accidentally getting ahead of the computer and possibly artificially inflating the error score.

Mercer, Harrell, Miller, Childs, and Rockers (1997), examined the performance of brain-injured versus healthy adults on the three versions of the Category Test. Forty-nine healthy adults and 45 brain-injured adults were tested using one of the three versions. No significant differences were found between the versions of the Category Test on total error score or on interaction between the Category Test version and group membership. This data supported the contention that the Category Test is robust to changes in the form of administration.

Common to all Category Test versions. Common to all of the adult versions is a 208 stimulus set, correct/incorrect feedback for each item, and a single response for each stimulus item. The 208 items are divided into 7 subtests, each with a common theme or principle running throughout. The first group consists of eight items and is usually easily performed, even by those with serious brain lesions (Reitan & Wolfson, 1985). It requires matching stimulus-item Roman numerals with Arabic numerals. Subtest 2 consists of 20 items and is a basic quantity test. For example, if there are two squares shown, the answer is two, if there are four circles shown, the answer is 4. Subtests 3, 4, 5, and 6 each contain 40 items and are based on a particular principle. These principles include odd stimulus, ordinal position, color (Spreen & Strauss, 1991), quadrant, or proportion (Reitan & Wolfson, 1985). Subtest 7 is not based on any given principle; it contains various items from the previous 6 subtests. This requires the subject to recall from previous viewings and give the correct response.

Scoring on the Category Test is based on the number of errors. Halstead (1947) suggested a cutoff score of 50. This means that a score of more than 50 indicated brain damage, while a score of 50 or less could be considered normal. Bigler, Steinman, and Newton (1982) indicated that the CAT was sensitive to age and that the mean error score rose dramatically for those over 50. Fromm-Auch and Yeudall (1983) only agreed with the 50-error cutoff for those under 40 years of age. In a sample of normal adults aged 41-64 (n=10), their median score was 53. Yeudall, Reddon, Gill, and Stefanyk (1987) reported age-stratified scores in 225 normal adults between 15 and 40 years of age. Group means were found to range from 27 to 35 errors on the CAT. Reitan and Wolfson (1985) presented severity ranges for the CAT scores as follows: 0-25, Perfectly Normal; 26-45, Normal; 46-65, Mildly Impaired; 65+, Seriously Impaired.

Origins of the Children's Category Test forms. The Category Test is an integral part of the Halstead-Reitan Neuropsychological Battery (Reitan, 1969; Reitan & Wolfson, 1992; Reitan & Wolfson, 1993) which was created by making significant modification of the tests originally included in the 1940 Halstead Battery. Reitan and Wolfson (1985) stated that the Halstead-Reitan Neuropsychological Battery has "probably been researched in more detail than any other set of neuropsychological tests" (p.10). The Halstead-Reitan Neuropsychological Battery (Reitan, 1969; Reitan & Wolfson, 1992; Reitan & Wolfson, 1993) actually encompasses three separate batteries (Boll, 1981). One battery was designed to assess the neuropsychological functioning of individuals 15 years and older. A second battery, the Halstead-Reitan Neuropsychological Test Battery for Older Children, is a revision of the adult battery and is used with children 9 to 14 years of age (Reitan & Davidson, 1974). The third battery is the Reitan-Indiana Neuropsychological Test Battery for Young Children, designed for children 5 to 8 years of age (Reitan, 1969). The children's batteries include modifications and a downward extension of the adult Halstead Reitan, as well as the addition of some supplementary measures not included in the adult version. The Halstead-Reitan Neuropsychological Test Battery for Older Children (Reitan & Davidson, 1974) contains a 168-item version of the Category Test, but the Reitan-Indiana Test Battery for Younger Children (Reitan, 1969) contains an 80-item Category Test.

The following is an example of how many of the shortened or altered forms of the original Category Test were created. A study was conducted by Reeder and Boll (1992) to determine if the Category Test contained in the Halstead-Reitan Neuropsychological Battery for Older Children could be shortened to improve its clinical usefulness while maintaining its psychometric properties. In this study, 12 shortened versions were generated from an administered original version. The short versions were based on those used in previous investigations, as well as other forms maintained the general format of the current version as an aid in retaining the basic characteristics of the test.

Total error scores for the 12 versions were generated and correlated with other neuropsychological variables. The other neuropsychological variables included the following: Wechsler Intelligence Scale for Children-Revised (Wechsler, 1974), the Wide Range Achievement Test-Revised (Jastak & Wilkinson, 1984), the Trail Making Test

(Reeder & Boll, 1992), the Tactual Performance Test (Reeder & Boll, 1992), the Finger Oscillation Test (Reeder & Boll, 1992), The Grooved Pegboard Test (Reeder & Boll, 1992), the Speech Sounds Perception Test (Reeder & Boll, 1992), the Seashore Rhythm Test (Reeder & Boll, 1992), and the Selective Reminding Test (Reeder & Boll, 1992). First the correlations were normalized (Fisher transformation procedure) and Z-tests were conducted to examine for correlational differences between the neuropsychological test scores and total Category Test scores of the original form with the 12 shortened forms. Results indicated that none of the Z-tests were statistically significant at $p < .05$. Zero-order correlations were then computed between the original form and each of the 12 shortened versions. The correlations were as follows: (1) .96; (2) .92; (3) .79; (4) .75; (5) .81; (6) .59; (7) .98; (8) .97; (9) .92; (10) .94; (11) .94; and (12) .93. Kuder-Richardson 21 internal consistency tests were also conducted to examine the homogeneity of variance of each of the abbreviated Category Test versions (Ghiselli, Campbell, & Zedeck, 1981).

Internal consistency test results for each of the forms were as follows: (1) 4.4; (2) 7.1; (3) 20.6; (4) 12.8; (5) 9.0; (6) 14.1; (7) 3.3; (8) 4.9; (9) 8.9; (10) 14.7; (11) 11.3; and (12) 6.3. Shortened forms 1, 2, and 12 were best overall on the combined criteria of (a) retaining the original format, (b) having the fewest items, (c) being highly correlated with the full-length Category Test, and (d) most closely resembling the full-length Category Test in relation to other instruments. A cross-validation study was then conducted between the three selected versions and a different subject group.

To perform the cross-validation, correlations were computed between long-version Category Test total error scores and the other neuropsychological test scores. Then shortened versions1 (administer first 20 items on subscales 1-6 = 108 total questions), 2 (administer first 20 items on subscales 1-5 = 88 total questions), and 12 (administer first 15 items scales 1-6 = 83 total questions) were calculated from administered long-form versions of the Category Test on the new subject data. Correlations were then computed between the total Category Test error scores on the shortened versions and the neuropsychological test scores. Correlations were normalized (Fisher transformation procedure) and Z-tests (Hays, 1981) were conducted examining for differences between the correlations of the three shortened versions with the neuropsychological test scores and the correlations of the long version with the neuropsychological test scores. No Z-test results were found to be significant at a $p < .05$ level of significance. Z-test results were then summed to determine how much each shortened version differed from the long version of the Category Test in the cross-validation sample. These summed Z-test results were 7.7, 12.3, and 12.8, for abbreviated versions 1, 2, and 12 respectively.

Lastly, Kuder-Richardson internal consistency tests were conducted to examine the homogeneity of variance of each of the abbreviated versions for the cross-validation group. Correlations were then computed between the original version and the three shortened versions. The correlations were: (1) .97, (2) .95, and (3) .95. Internal consistency ratings for each group were as follows: (1) .91, (2) .89, and (12) .87. Because version 12 had the fewest total items and were distributed across 6 subtests, it appeared to best represent a psychometrically and clinically sound alternative version to the long (168) Intermediate Category Test. This final version became the current Children's Category Test, Level 2 (for ages 9 to 16 years, 11 months).

The Children's Category Test

The Children's Category Test (CCT) was designed to assess non-verbal learning and memory, concept formation and problem-solving abilities in children 5 to16 years of age. The CCT also provides information on the child's ability to change problem-solving strategies, and to develop alternative solutions. The CCT assesses those abilities that involve problem-solving with novel material it is said to be a measure that assesses a child's "fluid" intelligence.

This theory of intelligence was developed by Raymond B. Cattell and John Horn (Cattell, 1963; Horn, 1985; Horn & Cattell, 1967). It includes two types of intelligence - fluid and crystallized. Fluid intelligence refers to essentially nonverbal, relatively culture-free mental efficiency, whereas crystallized intelligence refers to acquired skills and knowledge that are strongly dependent on their development on exposure to culture. Fluid intelligence involves adaptive and new learning capabilities and is related to mental processes, whereas crystallized intelligence involves overlearned and well established cognitive functions and is related to mental products and achievements. Sattler (1992) notes that fluid intelligence is more dependent on the physiological structures (for example, cortical and lower cortical regions) that support intellectual behavior than is crystallized intelligence. Fluid intelligence is also noted to be more sensitive to the effects of brain injury. Crystallized intelligence, which reflects cultural assimilation, is significantly influenced by formal and informal education factors. It is through the exercise of fluid intelligence that crystallized intelligence is said to develop.

Boll (1993) states, "the CCT directly assesses the cognitive processes required for successful academic achievement by measuring the child's ability to learn, to solve, and to develop, test, and modify hypotheses." Because the instrument is nonverbal in nature, the child's reasoning ability can be assessed independent of his or her expressive language skill level making the CCT is less educationally dependent than verbal reasoning measures. Thus, the CCT may be used to determine whether or not a child is able to perform learning-based processes despite the existence of learning disorders, verbal or motor deficits, emotional disturbance, or neurological deficits.

<u>CCT Level 1 and Level 2</u>. The CCT is made up of two separate levels: Level 1 is given to children 5 to 8 years of age and consists of 5 subtests and 80 items. Level 2 (as examined earlier) is administered to children ages 9 to 16 and consists of 6 subtests and 83 items. Both levels of the CCT follow the same theoretical tradition as Halstead's original Category Test, in that the stimuli presented follow several established rules or principles that vary across subtests. Level 2 subtests IV and V are the exception, as the principle is the same across both of these subtests. The child's task is to identify the idea or principle and apply it to succeeding items. Like the original Category Test, the child receives immediate feedback as to the correctness of each response.

On level 1, the child is shown a series of pictures that is intended to suggest a particular color. The chiid responds by pointing to or verbally identifying one of four colors printed on a response card. On level 2, the child is shown a series of pictures that is intended to suggest a particular number. The child responds by pointing to or verbally identifying the numeral 1, 2, 3, or 4 printed on a response card. On both levels, Subtest I (subtests are titled with roman numerals) can be employed as a practice test to determine if the child understands the task and is able to provide consistently appropriate responses. On the last subtest for both levels, the child is required to remember and apply the principles from the previous subtests. At each level, the CCT requires approximately 15 to 20 minutes to administer. Older children typically finish sooner than younger children taking the same level. Because concepts learned in each subtest are cumulative in the final subtest, the CCT must be administered in one session with no breaks between subtests.

CCT Standardization. Normative data on the CCT were derived from a standardization sample and based on an analysis of data gathered in March of 1988 by the U.S. Bureau of the Census and provided the basis for stratification along the following variables: age, gender, race/ethnicity, geographic region, and guardian education level. The standardization sample consists of 920 children in 12 age groups, ranging from 5 through 16 years of age. There were 80 children per age group in ages 5 through 12, and in age groups 13 to16 years, there were 70 children per age group. The CCT standardization sample consisted of 459 females and 461 males, with approximately equal numbers of both gender in each age group. For each age group, the proportion of Caucasians, African Americans, Hispanics, and other race/ethnic groups was based on March 1988 Census data and balanced within each age, gender, guardian education level, and geographic region. Information on guardian education level was obtained by asking parents to specify the highest grade completed by each parent (or guardian) living in the home. If both parents lived with the child, the average of the two education levels was used. If only one parent lived with the child, the education level of that person was used.

Concerning geographic region, the United States was divided into four major regions in accordance to the Census data: Northeast, North Central, South, and West. Children were selected for the standardization sample in accordance with the proportions of children living in each region.

The standardization sample was drawn from both the private and public school sectors. Students receiving mainstream special education services were not excluded. As a result, 7% of the standardization sample consisted of children who were classified as learning disabled, speech/language impaired, emotionally disturbed, or physically impaired. Only children who could understand English were tested. Lastly, 5.2% of the children were in gifted and talented education programs. Boll (1993) stated that the CCT standardization sample closely approximated the school-aged population of the United States as reported in the March 1988 Census data.

Reliability

Split-half reliability. Concerning the reliability of the CCT, one estimate of test score consistency is derived with the split-half reliability method. Boll (1993) made two equivalent halves of each of the two CCT levels representing parallel forms with approximately equal variances. Items were selected. Items were grouped into subgroups or "testlets" within each of the subtests, instead of a straight odd-even-item split across all subtests. These testlets were then grouped into two parallel forms. The scores on the half-tests were intercorrelated and corrected for the score on the full-length test by the Spearman-Brown formula. Averaged coefficients, based on Fisher's z transformation, were .88 for level 1 and .86 for level 2. The averaged coefficients, as well as the separate age-level values demonstrated a high degree of internal consistency.

Test-retest reliability. Test-retest reliability of the CCT was examined in a sample of 106 children (Boll, 1993). The sample was drawn from three age groups: 8 years of age (n = 35), 12 years of age (n = 40), and 16 years of age (n = 31). The interval between testings ranged from 10 days to 42 days, with a median retest interval of 28 days. The sample had the following composition: 49% female and 51% male, and 67% Caucasian children and 33% "minority" children.

The averaged test-retest reliability for the total error T score was .75. The test-retest coefficients for the three ages were as follows: Age 8 = .79, Age 12 = .70, and Age 16 = .75. The averaged coefficients indicated that the scores on the CCT possessed adequate reliability across short periods of time.

The T score changing between the first and second testings was consistent with previous research findings of practice effects. Although research on the test-retest reliability and practice effect on the CCT is sparse, some research has been conducted on the adult versions of the Category Tests (336 and 208 question versions). In 1966, Shaw found practice effects when examining the reliability of the Halstead Category Test (336), as did Seidenberg, O'Leary, Giordani, and Boll (1981) when they assessed the test-retest IQ changes of epilepsy patients on the CAT-208. Dodrill and Troupin (1975) examined the effects of repeated administrations of the Halstead Reitan Neuropsychological Battery (208) by administering the battery on four occasions at 6- to 12-month intervals to 17 epileptics with stable neurological dysfunction. Results indicated that if one used cutoff points as the sole basis of judgment with respect to normality of brain functions on the Halstead Category Test, one would judge twice as many to be normal by the forth administration as on the first.

Matarazzo, Weins, Matarazzo, and Goldstein (1974) computed test-retest reliabilities by administering the CAT-208 twice (20 weeks apart) to 29 normal males. Correlations between administrations were good ($r = .60$, p .001), indicating good reliability. As a comparison group, 16 60-year-old individuals with diffuse cerebrovascular disease showed test-retest correlations of .96 ($p < .001$), again showing very good reliability. DeFilippis and McCampbell (1979) found high correlations in a test-retest study on 30 normals tested three weeks apart ($r = .913$, $p < .001$) as well as for alcoholics ($r = .804$, $p < .001$) tested three weeks apart.

Rockers (1996) presented an experimental design that varied somewhat from traditional test-retest designs. For the first administration, Rockers varied the version of the Category Test that he administered: The versions included the Category Test (CAT-208) by Reitan and Wolfson (1985), the booklet version (BCT) by DeFilippis and McCampbell (1979), and the computerized version (CVCT-A) by Miller (1989). The second administration gave all three groups only the CVCT-A. The average time interval between test and retest was 30 days, with a minimum of four days and a maximum of 64 days. On initial administration, 11 subjects took the CAT, 9 took the BCT, and 10 took the CVCT-A. Rockers found that the error scores for the groups did not significantly differ from each other on the first administration and that post-test scores on the CVCT-A were significantly correlated with the three previously administered versions of the category test. Score differences from test to retest were not significant by group. This implied that, in terms of practice effects, all three versions were equal. In fact, whether the test was given at an interval of less than 26 days, or more, made no difference in the test scores. Rockers stated, " Thus, although scores may improve in normals on retest, the interval between is of relative unimportance."

MacInnes et al. (1981) in a test-retest cross-validation of CAT-208 and BCT, found low test-retest reliability for schizophrenic patients ($r = .521$, $p < .02$) and concluded that the Category Test is unreliable for schizophrenics. In other populations, test-retest correlations were high: organic brain syndrome ($n = 16$, $r = .785$, $p < .001$), personality disorders ($n = 9$, $r = .947$, $p < .01$).

Split-half reliability. Split-half and internal consistency also have minimal amounts of research cited but have been studied with the Adult version CAT-208. Kilpatrick (1970) computed split-half reliability by examining the protocols of 41 heterogeneous patients and found correlations between odd-item errors and even-item errors to be .97 ($p < .01$). Charter, Adkins, Alekoumbides, and Seccat (1987) also computed split-half reliability in a large unspecified sample ($n = 311$), corrected for age and education, and found high odd-even split-half reliability ($r = .95$). Moses (1985) evaluated internal consistency reliability in a sample of 285 diverse patients and found that the CAT-208 as a single score appears to have little measurement error (Cronbach's coefficient alpha = .96).

Validity

Construct-related reliability. CCT validity research can be placed into two traditional types: construct-related and criterion-related validity (Committee to Develop Standards for Educational and Psychological Testing, 1985). The CCT Level I and II subtests were designed as a measure of children's functioning in three general areas: nonverbal learning and memory, concept formation, and problem solving. Boll (1993) states, "from the Total Error Score, interpretations and inferences may be made about the latent or unobservable dimension of brain integrity or brain functioning." Correlations were examined between the CCT and the WISC-R Vocabulary Subtest, the Wechsler Intelligence Scale for Children-Third Edition (WISC-III), and the California Verbal Learning Test-Children's Version (CVLT-C).

All the children who were used in the standardization of the CCT were administered the WISC-R Vocabulary subtest (Boll, 1993). The correlations between the CCT raw score Total Error Score and the WISC-R Vocabulary subtest standard score were examined. The correlations ranged from -.23 to -.45. For Level I, the total correlation for all ages was .29 (p<.001) and for Level II the total correlation was -.33 (p<.001). These results indicated that the CCT consistently and significantly correlated with the WISC-R, even though both measure mental processes that require different types of abilities (i.e., verbal versus nonverbal abstract reasoning).

A clinical sample of 31 children, documented Attention-Deficit Hyperactivity Disorder (ADHD) were given both the CCT and WISC-III (Boll, 1993). The sample consisted of Caucasian males between the ages of 8 and 12 years of age. Modest correlations were found between the CCT total Error Score, the WISC-III Full Scale Intelligence Quotient (FSIQ) score ($\underline{r}$ = .22) , the Verbal Intelligence Quotient (VIQ), and the Performance Intelligence Quotient (PIQ) score. The CCT correlated more highly with the Performance subtests ($\underline{r}$ = .27) on the WISC-III than it did with the Verbal subtests ($\underline{r}$ = .14). This finding was expected given the nonverbal orientation of the CCT and is consistent with previous research done on the adult CAT-208 and the Wechsler Adult Intelligence Scale - Revised (Yeudall et al., 1987, Davidoff, Morris, Roth, & Bleiberg, 1985, Roth et al., 1989). In 1993, Mercer examined the relationship between the CAT total error score and the WAIS-R factors and scaled scores that assess short-term memory, attention and concentration, abstraction, psychomotor speed, and visuospatial skills on 49 healthy adults and 45 brain-injured adults. Of interest was the significant correlation ($\underline{r}$ = -.66, $\underline{p}$<.01) between the VSIQ and the Category Test total error scores.

Boll (1993) administered the entire CCT standardization sample (920 children) the California Verbal Learning Test - Children's Version (CVLT-C). The CVLT-C is a neuropsychological measure that assesses verbal learning and memory in children between the ages of 5 and 16. The CVLT-C index utilized for comparison to the CCT Total Error Score was the CVLT-C Total List A trials 1-5. This portion of the CVLT-C requires the child to spontaneously remember a list of 15 verbally presented words, with the benefit of word category cueing or cued recall. The CVLT-C Total List A index is the sum of all the words recalled over the five trials and provides a global measure of learning performance.

The correlations between the CCT raw score Total Error Score with the CVLT-C Total List A Index by age was as follows: 5 Years = -.23 ($\underline{p}$=.040), 6 years = -.16

(p=.153), 7 years = -.15 (p=.183), 8 years = -.27 (p=.015), 9 years = -.29 (p= .005), 10 years = -.20 (p=.082), 11 years = -.14 (p=.231), 12 years = -.37 (p=.0008), 13 years -.11 (p=.363), 14 years = -.22 (p=.070), 15 years = -.10 (p=.512), 16 years = -.27 (p=.026). The average correlation for both level I and level II was -.29 (p=.001). Boll (1993) stated, "although the degree of association between these two measures is fairly modest, it is within expected parameters and provides further evidence of the CCT's value as a measure of overall nonverbal ability."

The CCT has little data published specifically examining the functions it measures. The CAT-208 has More literature is published about the functions the CAT-208 measures, but there is considerable debate over what is actually measured (Bertram, Ables, & Snyder, 1990; Boyle, 1988; DeFilippis & McCampbell, 1991; Finlayson, Sullivan, & Alfano, 1986; Halstead, 1947; Reitan & Wolfson, 1985; Rothke, 1986; Sherrill, 1987; Taylor, Hunt, & Glaser, 1990). As mentioned previously, Halstead (1940) originally felt that the CAT measured "grouping" behavior-brain damaged subjects showed fewer and smaller grouping than did normals. In 1947, he stated that it measured integration and abstraction ability. Russell and Levy (1987) state that "abstraction" is a broad, encompassing term and should be broken into smaller elements. One element is extracting an element from a figure, whereas the other is the ability to shift principles or sets. Vanderploeg and Logan (1989) stated that in addition to the already mentioned set-shifting (between subsets) ability, there is another type of shifting required in the CAT-perceptual. This involves shifting within a subtest, for example, when the principle remains the same but the element type changes.

Rothke (1986) compared the CAT-208 with the Wisconsin Card Sorting Test and reported that the CAT was not so much a measure of set-shifting skill but more a measure of complex concept acquisition, visual memory and visuospatial reasoning, much like a general intellectual measure. Boyle (1988) also reported the CAT-208 as a measure of general intellectual ability.

Bertram, et al. (1990) reported that the CAT-208 could be viewed as a test of learning ability. Nussbaum and Bigler (1989) agreed that the CAT-208 measured learning along with skills of nonverbal reasoning and concept formation. Reitan and Davison (1974) reported that the CAT-208 not only a measured concept formation but also to apply organizing principles in a performance mode. Laatsch (1983) determined that the CAT-208 employed the development and application of rules and strategies. Golden, Osmon, Moses, and Berg (1981) went a step farther and expanded the idea of developing strategies into hypothesis formation and validation.

Choca, Laatsch, Wetzel, and Agresti (1997) stated, "It seems clear that the category test has great sensitivity but low specificity. In other words, individuals suffering from any kind of brain dysfunction, or having difficulty using their intellectual capacity effectively, are likely to perform poorly on this instrument." They reported that one may typically assume better scores on the CAT-208 to be the reflection of "an intact brain, reasonable intellectual abilities, maturity in cognitive development, and a capacity to think with concentration and efficacy."

Criterion-Related Evidence of Validity. Boll (1993) administered the original CCT and the current shortened version in a counterbalanced order to 110 children between the ages of 5 and 15, who were referred for a neuropsychological evaluation by the Neuropsychology Lab at the University of Alabama at Birmingham Hospitals. Referral diagnoses included traumatic brain injury, brain tumors, congenital cytomegalovirus, behavior disorders, and other neurological disorders (i.e., cerebral palsy, seizure disorder). Half of the children (n = 55; ages 5 to 8 years, 11 months) received both versions of level I. This group consisted of 33 males and 22 females who were predominantly Caucasian (34 Caucasian, 21 African American) with a mean WISC-R Full Scale IQ of 93.7 (SD = 18.9), Verbal IQ of 91.7 (SD = 18.3), and Performance IQ of 97.1 (SD = 18.4). The remaining children (n = 55; ages 9 years to 14 years, 11 months) were administered both versions of Level II. This group consisted of 32 males and 23 females who were also predominantly Caucasian (38 Caucasian, 17 African American), with a mean WISC-R Full Scale IQ of 89.9 (SD = 16.6), Verbal IQ of 89.9 (SD = 16.6), and Performance IQ of 91.6 (SD = 17.9).

All correlations among the subtests and total scores for the current CCT and the original versions of Level I were positive and significant with the exception of Subtest 1 (r = -.029). The results indicated that the two versions of Level I were strongly correlated with one another and are parallel forms. The correlations among the subtests and total scores for the current CCT and the original versions of Level II were significant, suggesting a strong positive relationship and equivalency between the two versions of Level II.

The initial presentation of Level 1 (Klonoff, Robinson, & Thompson, 1969) and Level II (Reed, Reitan, & Klove, 1965) demonstrated a consistently significant difference between neurologically impaired and normal children. A cross-validation of these tests by Reitan (1974) for Level I and Boll (1974) and Tsushima and Towne (1977) for both levels demonstrated that the CCT was a consistently valid indicator of normal and impaired brain functions in children ages 5 through 15.

Boll (1993) took children who were part of the CCT standardization sample and administered a group achievement assessment (n = 313). The subsample was evenly divided between males and females. The children were assessed by one of the following achievement instruments: the California Achievement Tests, Form E (CAT-E; CTB Macmillan/McGraw-Hill, 1988); the Iowa Tests of Basic Skills, Forms G and J (ITBS-G, ITBS-J; Hieronymus, Hoover, & Lindquist, 1986); the Metropolitan Achievement Test, Sixth Edition, Form L/M (MAT 6-L-M, Prescott, Barlow, Hogan, & Farr, 1985); and the Stanford Achievement Test, Seventh Edition Plus (SAT7; Forms E/F; SAT8: Gardner et al., 1987).

The CCT Total Error Score and composites of the Total Reading Score (consisting of reading comprehension and vocabulary) and the total Math Score (consisting of mathematical application and computation concepts) from the achievement batteries were compared. Correlations were calculated between normal-curve equivalent (NCE) scores on the composites and the CCT Total Error T score for three age groups (5 to 8, 9 to12, and 13 to16). Resulting correlations for Total Reading was as follows: ages 5 to 8 r = -.41, ages 9 to12 r = -.16, ages 13 to16 r = -.27, and Overall r = -.28. Resulting correlations for Total Math were: ages 5 to 8 r = -.47, ages 9 to12 r = -.30, ages 13 to16 r = -.27, and Overall r = -.34.

Correlations with School Grades

Boll (1993) obtained teacher assigned-grades for 313 children who were part of the CCT standardization sample. The grades were converted into a numeric system (50-100). Correlations were calculated between the CCT Total Error Score and the teacher-assigned grades for the subject areas of math, spelling, and reading. The following results were found for math: ages 5-8 had an $r = -.30$; ages 9-12 had an $r = -.25$; ages 13-16 had an $r = -.14$; and an overall $r = -.20$. For Spelling the results were: ages 5-8 the $r = -.27$; ages 9-12 the $r = -.44$; ages 13-16 the $r = -.59$; and overall the $r = -.38$. Lastly, Reading results indicated: ages 5-8 had an $r = -.15$; ages 9-11 had an $r = -.33$; ages 13-16 had an $r = -.08$; and overall the $r = -.20$.

Lastly, research on the CCT suggests there is a developmental age progression effect (Boll, 1993). On Level II, there is an increase in the number of errors within age groups on subtests II, III, and IV, which were intended to be the more difficult subtests, while subtest V measures the child's ability to recall previous concepts presented in earlier subtests. As the age increased, the total number of errors progressively decreased for all ages but 14 and 15. For ages 14, 15, and 16, the average number of errors was essentially the same, suggesting a leveling off of the error rate decline as a function of age.

Automation of the Category Tests

Many different types of psychological tests have been automated. Initially computer administration was introduced (Collins & Odell, 1986; Moerland, Aldencamp, & Alpherts, 1986; Morrison & Gates, 1988; Reitan & Davidson, 1974) but not far behind was test interpretation (Adams & Heaton; 1985, Kleinmutz, 1987). With the beginning of computers, the use, and possible misuse, of computer-based testing and interpretation have led to extensive debate. Neuropsychological tests have been no exception to varying viewpoints on the subject of automation; many have aired their disagreements (Adams & Heaton, 1985; Kleinmutz, 1987; Reitan & Davidson, 1974). In 1987, Adams and Heaton stated that the clinical deployment of computer-based interpretive programs for neuropsychology would be premature and potentially unethical.

Arguments for automated neuropsychological test administration include improved data collection (Collins & Odell, 1986; Rockers, 1996), standardized administration (Mooreland, et al., 1986; Morrison & Gates, 1988), and administrations independent of the clinician (Collins & Odell, 1986; Rockers, 1996). Arguments against neuropsychological automated test administration include the lack of standardization of computing environments (Moerland, et al., 1986) as well as the need for a "human" factor involvement (Reitan & Davidson, 1974).

Pro camp arguments concerning the automation of neuropsychological assessments include improved data collection and analysis the fact that a clinician need not be present during the assessment. Con camp arguments include that the current state of test data analysis leaves something to be desired, and that the clinician should always be present in an assessment situation. Adams and Heaton (1985) stated that the most appropriate role of the computer in test analysis is not to generate a final report but to provide a spreadsheet of calculations to aid the clinician in creating a total diagnostic package.

Regarding tests and test data, Adams and Heaton (1985) argued that the lack of agreement about diagnostic decision tools and failure to incorporate demographic variables (sex, age, education, etc.) that affect neurological performance into existing interpretation programs makes the trained clinician optimal; the clinician is superior to the computer. In 1987, Kleinmutz responded that the computer was "exquisitely equipped for clinical decision-making strategies because the machine is a noncomputational rather than a computational tool and because most expert and novice clinicians decisions can be articulated" (p.266). Thus, the problem was not in the nature or format of the data but in the need to formulate programs that capture a computer's processing capabilities.

In 1986 Collins and Odell administered a standard and computerized version of the Raven's Colored Progressive Matrices (Raven, 1965) to 16 aphasic individuals and found no significant differences between the two versions. Reaction times and error types were also reported to be no different than those of the standard administration. The Austin (or Milner) maze was examined by Morrison and Gates (1988) comparing the original (Walsh, 1978) and computerized version with sample of 32 male computer science students. Correlations between the two forms were not significantly different.

As mentioned previously, Beaumont (1975) was the first to investigate the validity of administering the CAT-208 via on-line computer. His findings were consistent with Goldstein and Shelly (1972) who had examined the traditional administration of the CAT-208. In 1989, Miller developed a computerized version of the Category Test for older children (CV-CAT) for the Macintosh computer. When the Intermediate version of the BCT and the Computerized Version of the Category Test-208 (CV-CAT) were compared, no significant differences were found for total error or subtest errors. Notably different was the completion time for the CV-CAT versus the BCT. The completion time for the

CV-CAT was significantly longer than for the BCT. Miller (1989) suggests that this difference may have resulted in the time required for the individual to acclimate to the mouse pointer in the computerized version.

As mentioned earlier, Rockers (1996) examined the test-retest correlation of the Computerized Version of the Category Test: Adult Version (CVCT:A) and found a less-than-expected correlation due to problems involving the subject clicking the mouse twice in order to respond to a computer question instead of once, thus accidentally getting ahead of the computer and possibly artificially inflating the error score. Also mentioned earlier, Mercer, Harrell, Miller, Childs, and Rockers (1996), examined the performance of brain-injured versus healthy adults on the three versions of the CVCT:A and found no significant differences between the Category Test version and group membership.

<u>The Computerized Version of the Category Test-Young Children's Version (CVCT-YC).</u>

In 1994, Miller developed a computerized version for level one of the CCT for the Macintosh computer: the Computerized Version of the Children's Category Test-Young Children's Version (CVCT-YC). To date, many of the psychometric properties of the CVCT-YC have yet to be examined. This computerized version provides a permanent record of responses as well as a record of item response time and test administration time.

Existing research indicates that the computerized version is a useful alternative to the traditional booklet format (despite aired concerns about assessment automation, the ease of administration, data collection, and analysis). Test-retest reliability coefficients range from .60 for normals to .96 for neurologically impaired individuals (Matarazzo et al., 1974; Matarazzo et al., 1976; Sarazin & Spreen, 1986). Split-half reliability coefficients range from .94 (Kilpatrick, 1970) to .97 (Charter et al., 1987). Also, little doubt exists that all adult versions of the Category Test are sensitive measures of neuropsychological impairment (Reitan & Davidson, 1974; Dodrill & Troupin, 1975; Golden et al., 1981; Reitan & Wolfson, 1985; Boyle, 1988; Spreen & Strauss, 1991).

The CVCT-YC, has significant promise concerning the neurological assessment of children. The CVCT-YC, like the CCT was designed to measure abilities related to reasoning, abstraction, logical analysis, nonverbal learning, memory, concept formation, and general problem-solving abilities in children (Boll, 1993; Reitan & Wolfson, 1992). The CVCT-YC, like the CAT-208, should be a robust discriminator in the assessment of brain-damaged versus normal functioning individuals.

However, many questions about the CVCT-YC need to be further examined. Reliability: How consistent is this instrument over time? Validity: How will measures of fluid reasoning from an instrument that is already psychometrically established compare to results from the CVCT-YC? Will results on the CVCT-YC be affected by gender, ethnicity, age? Studies have shown possible practice effects between the initial and second administration with the CAT-208; does the CVCT-YC exhibit the same? Even with practice effects, could alternative forms of the CVCT-YC be developed and tested, which may then allow it to be helpful in providing data concerning monitoring the benefits of remediation and medication interventions in young children? Is this form of assessment capable of all that it is thought to be? We do know that the CVCT-YC requires minimal space, provides a permanent record of responses, as well as a record of item response time and test administration time, and potentially eliminates examiner error in response scoring and administration. The CVCT-YC may be one of the answers to one of the major questions concerning child neuropsychology.

At the beginning of the chapter the following question was posed: Which tests or diagnostic approaches are appropriate for use with children (Hartlage & Long, 1997)? It was mentioned that, especially in the schools, something must be found to screen children for potential brain damage that is quick and economical enough to administer as a type of screening device, yet can also be an assessment that could indicate possible neurological problems. At this point a practitioner could justify a more lengthy and in-depth neurological examination of the child. The CVCT-YC may be the assessment tool that can fulfill the current neurological assessment demands discussed earlier.

However, before the CVCT-YC can be proclaimed an answer to anything, it must exhibit itself to be a valid and reliable method of assessment. It must provide a firm foundation of data supporting the basic elements of test reliability and validity. To date, very little has been done concerning the psychometric properties of the CVCT. The current research addresses this very problem.

Purpose

The properties of the Computerized Version of the Children's Category Test-Young Children's Version (CVCT-YC) are currently being researched by several individuals. The present study will examine the test-retest reliability and concurrent validity of the CVCT-YC in a normal student population. In addition, scores for gender, age, hand dominance, ethnicity, income, education levels of parents, guardian marital status, and with whom the child resides will be examined for the possible relationship between the CVCT-YC and Concept Formation scores from the Woodcock-Johnson Tests of Cognitive Ability.

Based on a review of the literature, the following hypotheses appear warranted:

1. Test-retest total error scores of the CVCT-YC will be significantly correlated with the initial assessment scores administered three weeks prior.
2. Error scores on each subtest will significantly correlate with error scores from a readministration three weeks later.
3. Total error scores from the initial administration of the CVCT-YC will be negatively correlated with standard scores on the subtest Concept Formation from the Woodcock-Johnson Tests of Cognitive Ability.
4. There will be no significant relationship between age and CVCT-YC and WJ-R Concept Formation scores.
5. There will be no significant relationship between gender and CVCT-YC and WJ-R Concept Formation scores.
6. There will be no significant relationship between ethnicity and CVCT-YC and WJ-R Concept Formation scores.
7. There will be no significant relationship between hand dominance and CVCT-YC and WJ-R Concept Formation scores.
8. There will be no significant relationship between income and CVCT-YC and WJ-R Concept Formation scores.
9. There will be no significant relationship between mother's level of education and CVCT-YC and WJ-R Concept Formation scores.

10. There will be no significant relationship between father's level of education and CVCT-YC and WJ-R Concept Formation scores.
11. There will be no significant relationship between marital status of the primary guardian and CVCT-YC and WJ-R Concept Formation scores.
12. There will be no significant relationship between whom the child resides with and CVCT-YC and WJ-R Concept Formation scores.

CHAPTER II

Method

This chapter is divided into five sections. In the first section, Participants, a description of the population from which the sample was selected and a description of the subject characteristics within the sample are given. The second section, Instruments, discusses The Computerized Version of the Category Test-Young Children's Version and the Concept Formation subtest from the Woodcock-Johnson Tests of Cognitive Ability. The third section, Procedures, discusses the specific procedures followed in the study. The fourth section, Data Analysis, explains what statistical procedures were used.

Participants

The number of participants in this study was originally 75 elementary school children between the ages of 5 and 8, but due to attrition 56 subjects completed the study: 24 females 42.9% and 32 males 57.1%. The subjects were drawn from several YMCA after-school programs in northern Texas and the mean age was 6.78 years (See Table 1).

Table 1.

<u>Subject Age Distribution and Percentages</u>

Age	n	Percent of Sample
5-year-olds	07	12.5%
6-year-olds	14	25.0%
7-year-olds	20	35.7%
8-year-olds	15	26.8%

Attempts were made to follow the 1990 U.S. Census Data concerning appropriate representation of ethnicity for children ages 5 to 9 (See Table 2).

Table 2.

<u>1990 U.S. Census Ethnic Percentages Compared to Sample Data Percentages</u>

Ethnicity	Census Population	Sample Population
White Anglo American	71.3%	60.71%
African American	12.1%	10.71%
Asian American	2.9%	8.93%
Hispanic American	9.0%	7.15%
Native American	.8%	5.36%
Other	3.9%	7.14%

The majority of the subjects came from upper income families. Mean family income was collected and listed in Table 3.

Table 3.

Mean Family Incomes Per Year

Income Group	n	Percent of Sample Population
15,000 - 29,999	03	5.4%
30,000 - 44,999	07	12.5%
45,000 - 59,000	02	3.4%
60,000 - 74,999	07	12.5%
75,000 - 84,999	12	21.4%
85,000 - 99,999	01	1.79%
Over 100,000	16	28.6%
Not Known	08	14.3%

Due to possible performance problems in controlling the computer mouse, data was collected on subject handiness. Of the total subjects, 80.4% (n = 45) were right-handed and 19.7% (n = 11) were left-handed. The majority of the subjects lived with their biological parents (64.3%, n = 36), while adoptive parents (3.6%, n = 2), step-parents (5.4.%, n = 3), and single parents (16%, n = 9) were in the minority. The majority of subject primary caregivers were married (69.6%, n = 39) as opposed to single (10.7%, n =

6), or divorced (8.9%, n = 5). (10.7 %, n = 6 did not complete this question) Table 4 lists a breakdown of parent's highest levels of education.

Table 4.

Parent Highest Education Level

	Mother's		Father's	
Highest Level of Education	Percent	n	Percent	n
High School or GED	17.9%	10	8.9%	05
2yr Associate's Degree	19.6%	11	21.4%	12
4yr Bachelor's Degree	35.7%	20	41.0%	23
Master's Degree	12.5%	07	10.7%	06
Doctoral Degree	1.7%	01	1.7%	01
Other	12.5%	07	16.7%	09

All subjects were treated in accordance with the "Ethical Principles of Psychologists and Code of Conduct" (American Psychological Association, 1992).

Instruments

The Computerized Version of the Category Test-Young Children's Version (CVCT-YC) (Miller, 1994) and the Concept Formation Subtest from the Woodcock-Johnson Tests of Cognitive Ability (Woodcock, 1977) were the two primary instruments administered in this study. The purpose of the CVCT-YC is to measure a child's problem-solving skills and cognitive flexibility. The CVCT-YC required the subjects to look at a stimulus item presented on a computer screen, then decide what color (red, blue, or green) best described the picture. For example, if a red circle was shown to the subject, then the correct answer would be red. After responding, the subject received feedback that indicated if the answer was correct or incorrect. The subject was asked to look at 80 images individually and categorize them. The 80 items on the CVCT-YC were divided into four subtests. On the first four subtests there was one rule or principle that is the same across all items. The last subtest is a memory test, which asked the subject to recall correct answers from previous sections. Each test administration took approximately 15 minutes. Stimulus items were presented to the child via a Macintosh computer. The subject indicated a response by positioning a mouse and "clicking" on the correct response.

Concept Formation is a subtest from the Supplemental Battery of the Woodcock-Johnson Tests of Cognitive Ability; the test primarily measures reasoning or fluid intelligence. The Concept Formation Test exhibits split-half reliability coefficients of .953 for 6-year-olds ($\underline{n}$= 305, $\underline{\text{SEM}}$ ($\underline{\text{W}}$) = 3.6) and .934 for 9-year-olds ($\underline{n}$ = 310, $\underline{\text{SEM}}$ ($\underline{\text{W}}$) = 4.0). The mean reliability coefficient for ages 6-9 was .934 ($\underline{\text{N}}$ = 1840, $\underline{\text{SEM}}$ ($\underline{\text{W}}$) = 4.5). Concept Formation is a controlled learning task that, like the CVCT-YC, involves categorical reasoning which does not have a memory component. The tasks require reasoning based on principles of formal logic. The subjects were required to look at a set of objects either presented inside or outside of a box. The subject is asked to identify the rule for an object to be placed inside the box. Like the CVCT-YC, the subject was given feedback regarding the correctness of each response. The subject was asked between 6 and 35 questions depending on their ability. The dependent measure on the Concept Formation Test was the number of correct categorizations, unlike the CVCT-YC, that scored based on the number of errors.

Every parent or primary caregiver was given a questionnaire to fill out along with the consent form. The questionnaire was not required to be completed in order for their child to participate in the research. The questionnaire asked questions concerning gender, ethnicity, income, parent education level, who the subject lived with, and parent or guardian marital status (See Appendix A).

Procedure

Upon receiving consent from the YMCA director, the YMCA board of directors, and each site's child-care director, a letter of consent and questionnaire was personally given to each participating parent or primary guardian (See Appendix B). Most parents completed the consent form and questionnaire at the time they were presented with the research information verbally. All parents were asked if their children had any formal diagnosis of Attention Deficit Disorder, Brain Impairment, Learning Disability, Mental Retardation, or Emotional Disturbance. The children who did have any of the above-mentioned diagnoses were excluded from the study.

Each subject participated in two test sessions approximately three weeks apart (all retest dates fell within one or two days of the three-week mark). In both sessions, the assessor began by making sure the subject appeared at ease and then directed the subject's attention to the computer program. The assessor made sure that the subject understood the test directions and how to operate the computer mouse properly. Each subject was allowed to practice moving the mouse prior the actual CVCT-YC administration. The subjects were monitored by the examiner at all points during the assessment sessions. Specific directions were stated via the computer to the subject, but the examiner did occasionally repeat or elaborate upon the instructions if the subject appeared confused or did not understand the task. The subject was not given the underlying principle of the subtest. The child was not asked to verbalize the reason for a correct response, because that would have shifted the task from nonverbal reasoning and memory to verbal reasoning and memory.

After the initial administration of the CVCT-YC, each subject was administered the Concept Formation Subtest from the Woodcock-Johnson Tests of Cognitive Ability. All subjects began with the first item (Basal), and the ceilings were established based on cutoff points throughout the administration. Directions were read from the administration booklet, and instructions on what can be said and cannot were very clear. The rules for administration are very similar to those presented for the CVCT-YC. The subject could be administered a minimum of six questions through a maximum of 35 questions. As mentioned earlier, scoring was based on the number of correct responses then converted into a standard score. The Concept Formation was only administered to the subject once during the initial assessment session after the CVCT-YC was completed by the subject.

Approximately three weeks later, the subject was readministered the CVCT-YC following the same administration procedures mentioned above. All questionnaires, CVCT-YC score sheets and Concept Formation protocols were given numbers at the time of administration based on date, sequence, and location. No last names were used on any forms, with the exception of parent signatures. After the retests were administered, each subject's data was renumbered at random so that all of the data had no school, or center related sequencing and then entered into the data bank for statistical analysis.

CHAPTER III

Results

Hypothesis 1: Post-test total error scores of the CVCT-YC will be significantly correlated with the initial assessment scores administered three weeks prior.

Hypothesis 1 Analyses: A Pearson product-moment correlation was computed. Results presented a reliability coefficient of .90 between the post-test total error scores and the initial assessment error scores on the CVCT-YC ($p<.001$).

Hypothesis 2: Error scores on each subtest will significantly correlate with error scores from a readministration three weeks later.

Hypothesis 2 Analyses: A Pearson product-moment correlation was computed. Results are presented in Table 5.

Table 5.

Reliability Coefficients for the Total Sample by CVCT-YC Test Sections and Total Test

	Test		Retest		
Test Section	Mean	SD	Mean	SD	r
Section I	0.21	0.53	0.18	0.61	.33
Section II	3.90	3.61	3.07	2.80	.81
Section III	3.07	3.18	1.60	2.70	.77
Section IV	3.68	5.09	2.23	4.20	.78
Section I-V	10.86	9.49	7.10	7.21	.89

Table 5 (Cont.).

Reliability Coefficients for the Total Sample by CVCT-YC Test Sections and Total Test

	Test		Retest		
Test Section	Mean	SD	Mean	SD	r
Section V	1.37	1.54	0.87	1.45	.78
Total Test	12.21	10.73	7.98	8.17	.90

Hypothesis 3: Total error scores from the initial administration of the CVCT-YC will be negatively correlated with standard scores on the subtest Concept Formation from the Woodcock-Johnson Tests of Cognitive Ability.

Hypothesis 3 Analyses: A Pearson product-moment correlation was computed and the results presented a correlation coefficient of -.441 ($p < .001$).

Hypothesis 4: There will be no significant relationship between age and CVCT-YC and WJ-R Concept Formation standard scores.

Hypothesis 4 Analyses: A Hotelling T Square Test was used to see if the groups differed on the two dependent variables combined. The Hotellings T Square Test indicated no significant age effect ($F(6, 100) = 1.29$, $p = .27$). Mean standard scores for the Concept Formation Test are listed in Table 6.

Table 6.

Based on Age, Mean Standard Scores for the Concept Formation Test.

Age	MSS	SD
5-year-olds	107	15
6-year-olds	100	13
7-year-olds	105	13
8-year-olds	111	13

Reliability coefficients were also calculated based only on the CVCT-YC test sections and total test scores based on age (see Table 7). Results indicated strong reliability coefficients between test and retest means with the exception of Section I for all of the age groups and Section III for the 8-year-olds. In addition, the 8-year-olds did not score any errors on Section I of their retest, making a reliability coefficient unavailable. Fisher's Z-transformations were computed from reliability results and examined for significant differences between the age groups (see Table 8).

Table 7.

Reliability Coefficients for the CVCT-YC Test Sections and Total Test Scores Based on 5-and 6-year-olds.

	5-Year-Olds					6-Year-Olds				
	Test		Retest			Test		Retest		
Test Section	Mean	SD	Mean	SD	r	Mean	SD	Mean	SD	r
Section I	1.0	1.0	0.6	1.1	.59	.14	0.4	0.1	0.3	-.11
Section II	5.0	3.2	4.1	3.8	.78	5.7	4.1	4.3	3.6	.91
Section III	4.4	4.5	1.6	1.9	.81	3.8	4.6	2.1	3.9	.86
Section IV	4.9	6.3	5.0	6.6	.99	3.7	5.2	2.7	4.5	.90
Section I-V	15.3	10.3	11.3	7.8	.79	13.4	11.3	9.2	9.5	.96
Section V	1.7	2.0	1.1	1.3	.70	2.2	2.0	1.4	1.5	.82
Total Test	17.0	12.2	12.4	8.8	.84	15.6	13.0	10.6	11.0	.95

Table 7. (continued)

Reliability Coefficients for the CVCT-YC Test Sections and Total Test Scores Based on 7-and 8-year olds.

	7-Year-Olds					8-Year-Olds				
	Test		Retest			Test		Retest		
Test Section	Mean	SD	Mean	SD	r	Mean	SD	Mean	SD	r
Section I	.1	.3	.3	.7	-.12	.1	.3	.0	.0	1.0
Section II	3.2	2.6	2.7	2.1	.71	2.5	2.8	2.0	2.0	.67
Section III	2.6	4.0	1.7	3.0	.80	2.4	2.2	1.0	1.0	.25
Section IV	3.8	5.4	1.2	2.9	.60	3.0	4.2	2.0	3.9	.83
Section I-V	9.7	10.0	5.8	6.0	.90	8.0	5.2	4.9	4.8	.69
Section V	1.0	1.2	0.6	0.8	.61	0.9	0.9	0.6	0.7	.61
Total Test	10.7	10.8	6.4	6.8	.92	8.9	5.8	5.5	5.2	.72

Table 8.

Fisher's Z-transformation comparisons Based on Age

Group	Z	
5-year-olds vs. 6-year-olds	1.05	Not Significant
5-year-olds vs. 7-year-olds	- 0.72	Not Significant
5-year-olds vs. 8-year-olds	0.54	Not Significant
6-year-olds vs. 7-year-olds	0.54	Not Significant
6-year-olds vs. 8-year-olds	2.21	Significant
7-year-olds vs. 8-year-olds	1.90	Not Significant

Fisher's Z-transformation scores indicated a significant difference between the 6-year-old and the 8-year-old groups' Total Test Scores ($p < .01$). Fisher's Z-transformation scores were then calculated for each CVC-YC Section and found the following to be significant: Section I ($p < .01$), Section III ($p < .05$), and Section I-IV ($p < .01$).

Hypothesis 5: There will be no significant relationship between gender and CVCT-YC and WJ-R Concept Formation scores.

Hypothesis 5 Analyses: A Hotelling T Square Test was used to see if the groups differed on the two dependent variables combined. The Hotelling T Square Test indicated no significant gender effect ($F = (2, 53) = 2.54$, $p = .09$). Reliability coefficients for the CVCT-YC test sections and total test scores based on gender were calculated and presented in Table 9.

Table 9.

Reliability Coefficients for the CVCT-YC Test Sections and Total Test Scores Based on Gender of the Subject.

	Males					Females				
	Test		Retest			Test		Retest		
Test Section	Mean	SD	Mean	SD	r	Mean	SD	Mean	SD	r
Section I	.2	.5	.2	.6	.09	.3	.5	.2	.6	.64
Section II	4.3	3.9	3.2	3.1	.89	3.4	2.6	2.9	2.5	.63
Section III	2.6	4.0	1.7	3.0	.80	2.4	2.2	1.0	1.0	.25
Section IV	3.8	5.4	1.2	2.9	.60	3.0	4.2	2.0	3.9	.83
Section I-V	9.7	10.0	5.8	6.0	.90	8.0	5.2	4.9	4.8	.69
Section V	1.0	1.2	.6	.8	.61	.9	.9	.6	.7	.61
Total Test	10.7	10.8	6.4	6.8	.92	8.9	5.8	5.5	5.2	.72

Fisher's Z-transformations were computed based on gender. Results indicated a significant difference between CVCT-YC total reliability coefficients and gender ($p < .05$). Males exhibited significantly higher reliability coefficient than the females. Significant differences based on gender were also found on CVCT-YC Sections I ($p < .05$), II ($p < .05$), and III ($p < .01$).

Hypothesis 6: There will be no significant relationship between ethnicity and CVCT-YC and WJ-R Concept Formation scores.

Hypothesis 6 Analyses: A Hotelling T Square Test was used to see if the groups differed on the two dependent variables combined. The Hotelling T Square Test indicated no significant ethnicity effect (F (10, 96) = .752, p = .67). Reliability coefficients were calculated for the CVCT-YC Test Sections and Total Test Scores based on the ethnicity of the subject (see Table 10).

Table 10.

Reliability Coefficients for the CVCT-YC Test Sections and Total Test Scores Based on Ethnicity of the Subject.

	African American					Asian American				
	Test		Retest			Test		Retest		
Test Section	Mean	SD	Mean	SD	r	Mean	SD	Mean	SD	r
Section I	.7	1.0	.5	1.2	.63	.0	.0	.0	.0	.0
Section II	4.2	2.1	4.0	3.1	.62	3.4	2.8	3.2	1.7	.56
Section III	4.0	5.1	1.3	1.9	.97	2.0	1.9	.1	.8	.79
Section IV	5.5	7.8	4.6	7.0	.98	5.8	4.1	1.4	2.6	.03
Section I-V	14.3	13.6	10.5	7.9	.93	11.2	3.6	5.4	4.9	.86
Section V	2.2	1.9	1.5	1.4	.71	1.2	1.3	.6	1.3	.34
Total Test	16.5	15.3	12.0	8.9	.96	12.4	4.0	6.0	6.2	.84

Table 10. (continued)

Reliability Coefficients for the CVCT-YC Test Sections and Total Test Scores Based on Ethnicity of the Subject.

	Hispanic American					Native American				
	Test		Retest			Test		Retest		
Test Section	Mean	SD	Mean	SD	r	Mean	SD	Mean	SD	r
Section I	.0	.0	.0	.0	.00	1.0	1.0	.3	.6	.87
Section II	2.5	2.3	3.0	2.1	.71	10.3	4.5	9.3	2.3	.89
Section III	.5	.6	.8	.5	-.57	6.0	7.0	2.0	3.5	1.0
Section IV	3.0	4.6	.5	.6	.61	3.0	3.4	2.6	2.8	1.0
Section I-V	6.0	4.9	4.2	2.3	.71	20.3	14.5	14.3	8.3	1.0
Section V	1.7	1.0	.5	1.0	.87	3.0	3.4	2.0	1.7	1.0
Total Test	7.7	5.6	4.7	2.9	.72	23.3	18.0	16.3	10.1	1.0

Table 10. (continued)

Reliability Coefficients for the CVCT-YC Test Sections and Total Test Scores Based on Ethnicity of the Subject.

	White Anglo American				
	Test		Retest		
Test Section	Mean	SD	Mean	SD	r
Section I	.1	.4	.1	.3	-.13
Section II	3.5	3.2	2.3	2.4	.82
Section III	2.9	3.3	1.5	2.4	.71
Section IV	2.8	4.6	2.0	4.0	.88
Section I-IV	9.4	8.3	5.9	6.9	.86
Section V	1.0	1.19	.7	1.0	.83
Total Test	10.3	9.11	6.6	7.8	.88

Fisher's Z-transformations were computed based on ethnicity, and no significant differences were found between CVCT-YC total reliability coefficients and ethnicity.

Hypotheses 7, 8, 9, 10, 11, and 12 Analyses: There will be no significant relationship between the following and CVCT-YC and WJ-R Concept Formation scores:

subject hand dominance	income
mother's level of education	father's level of education
marital status of primary guardian	with whom the subject lives with

Hypotheses 7, 8, 9, 10, 11, and 12 Analyses: A Hotelling T Square Test was administered for each hypothesized possible effect to see if a difference existed on the two dependent variables combined. The Hotelling T Square Test indicated no significant affects (See Table 11). No further analysis was conducted on these areas of interest due to very small subject samples.

Table 11.

Hotelling T Square Test Results for Hypotheses 7 through 10

Hypothesis	Effect	F	df	p
Hypothesis 7	Hand Dominance	.58	4, 102	.67
Hypothesis 8	Income	1.07	14, 92	.39
Hypothesis 9	Mother's Education	1.58	10, 96	.13
Hypothesis 10	Father's Education	.81	10, 96	.62
Hypothesis 11	Caregiver Marital	1.10	6, 100	.36
Hypothesis 12	Child Lives With	1.79	10, 96	.07

*Note. All multivariate test statistics were calculated only with specific subject response, no value was given to subject non-response items.

CHAPTER IV

Discussion

The primary focus of the current research involved the reliability and validity of the Computerized Version of the Category Test - Young Children's Version (CVCT-YC). Statistical results indicate that the CVCT-YC is indeed statistically stable over time, specifically over three-week periods. The CVCT-YC was also found to correlate significantly with a psychometrically established measure of fluid reasoning: the Concept Formation Test, which originates as a subtest from the Woodcock-Johnson Tests of Cognitive Ability. This research also examined for possible significant relationships between CVCT-YC and Concept Formation scores and particular populations. This research examined the relationship between CVCT-YC and Concept Formation scores and age, gender, hand dominance, ethnicity, income, mother's highest level of education, father's highest level of education, primary caregiver's marital status, and with whom the subject lives. None of the populations examined were found to exhibit a significant relationship with assessment scores.

However, when a Fisher's Z-transformation was calculated on scores from the reliability coefficients for the CVCT-YC test sections and total test scores based on age, significant comparisons were found between the 6-year-olds and the 8-year-olds. Upon closer investigation it was found that specifically test sections I, III, and I-IV were found to be significant. On Section I, the 6-year-olds exhibited a negative reliability coefficient, whereas the 8-year-olds missed no questions. However, on sections III and IV the 6-year-olds exhibited stronger reliability coefficients. Why these two groups differed is difficult to pinpoint. Why all of the age groups failed to display a strong reliability coefficient between Section I test and retest scores could also be due to many things but further testing created to examine these specific areas is definitely warranted. However, through observing the subjects during testing, several possible factors were noted.

Initially, during many of the assessment sessions, the subjects were very anxious and/or apprehensive about the computer, the mouse, and their performance, but by section II most subjects appeared more comfortable and less anxious. Usually by the time of the retest, subjects appeared to be less anxious and more relaxed through the entire assessment. Negative correlation coefficients could be explained by some subjects actually becoming too comfortable with their response actions and missing more than they had initially. In other cases, especially with the older children, the subjects would become especially nervous if they began to miss questions and appeared embarrassed that the examiner could hear that they were not performing correctly. Often, beginning on Section III, several of the subjects appeared to become so anxious about their performance that upon missing just one item they would continue to miss questions until they settled down enough to concentrate on the task at hand. It would be very interesting to measure a subject's level of anxiety during testing with the CVCT-YC and compare it to their performance.

It is also suspected that the order in which the CVCT-YC and WJ-R Concept Formation were given during the initial assessment session, the time of day in which all research was collected, and the time of year in which the research collection took place, could all have had an affect on assessment results. All subjects were first administered the CVCT-YC and then given the WJ-R Concept Formation Test. Most of the subjects appeared very curious and excited about the computer. The CVCT-YC appeared to keep the subject's attention better than the Concept Formation Test (that is administered from a booklet). The younger ones especially appeared to tire of the testing process very quickly and perhaps did not put forth their continuing best effort. All assessments were conducted during after-school hours, which ranged from 2:30pm until 6:30pm. Again, many subjects may have been fatigued at that hour of the day. Lastly, the research collection process took place during the last month of school, in which many children may have been anxious for school to end and excited about end-of-year events. Continuing research needs to be conducted during various times of the day and year.

The second area that was found to be significant, when examining CVCT-YC total reliability coefficients, was gender. Again, it is difficult to state precisely why males and females performed differently. Is it a question of biological differences? Do females tend to remember things longer, recall better, or become less frustrated when they make an error? It would interesting to see if specific age and gender would affect scores. On average, would 6-year-old girls perform significantly worse than 8-year-old boys? How would 6-year-old boys perform compared to the 6-year-old females?

Lastly it is important to note that on the questionnaire which asked parents about their child's particular ethnicity, it is suspected that several parents marked their children to be of Native American status instead of checking White Anglo American. In many cases, it is likely that their child may be a native of America but not of direct Native American descent.

Concerning the questions stated at the end of Chapter 1, several have already been addressed. Reliability: How consistent is this instrument over time? This research has found the CYCT-YC to be significantly consistent over a period of three weeks. Validity: How will measures of fluid reasoning from an instrument that is already psychometrically established compare to results from the CVCT-YC? Research correlation coefficients from the CVCT-YC and the Concept Formation Test (a subtest from the WJ-R Tests Of Cognitive Abilities) were found to be significantly correlated. Will results on the CVCT-YC be affected by age, gender, hand dominance, ethnicity, income, mother's highest level of education, father's highest level of education, primary caregiver's marital status, and with whom the subject lives? Possibly, this research indicated significant areas of interest concerning age and gender but more research needs to be collected due to the current sample size available for analysis. The subject pool was made up of primarily upper-income subjects from one specific area of the country which limits subject pool generalization.

Studies have shown possible practice effects between the initial and second administration with the CAT-208 (the adult version of the Category Test); does the CVCT-YC exhibit the same? No definite evidence of practice effects was found to significantly effect the test-retest correlation coefficients in this research. If practice effects had existed they, did not appear to affect the scores enough to make them not significant.

Could alternative forms of the CVCT-YC be developed and tested which may allow it to be helpful in providing data concerning monitoring the benefits of remediation and medication interventions in young children? This research has found the reliability of the CVCT-YC to be a stable measure over time, and it does not appear to be significantly affected by many extraneous variables. However, more research needs to be focused on the areas that this research found to have a possible effect on test results. Once questions concerning age and gender can be answered, or at least predicted, and accommodated into the known psychometric qualities of this test the tradition may continue - the tradition which began with the original Category Test that has been modified many times, creating alternative reliable and valid forms of assessment.

Is this form of assessment capable of all that it is thought to be? The current research proves to be just one piece of a picture that is being created to exhibit the strength of the possibilities and assessment potential of the CVCT-YC. Much more research needs to be completed to confirm just how helpful and capable the CVCT-YC actually is. We do know that the CVCT-YC requires minimal space, provides a permanent record of responses, as well as a record of item response time and test administration time, and potentially eliminates examiner error in response scoring and administration.

Lastly, at the beginning of the chapter the following question was posed: Which tests or diagnostic approaches are appropriate for use with children (Hartlage & Long, 1997)? It was mentioned that, especially in the schools, something must be found to screen children for potential brain damage that is quick and economical enough to administer as a type of screening device, yet is an assessment that could indicate possible neurological problems, at which point a practitioner could justify a more lengthy and in-depth neurological examination of the child. The CVCT-YC may be the assessment tool that can fulfill the current neurological assessment demands discussed earlier. However, before the CVCT-YC can be proclaimed an answer to anything, it must exhibit itself to be a valid and reliable method of assessment. It must provide a firm foundation of data supporting the basic elements of test reliability and validity. To date, very little has been done concerning the psychometric properties of the CVCT-YC. The purpose of the current research is to provide a beginning step in addressing this problem.

References

Adams, K. M., & Heaton, R. K. (1985). Automated interpretations of neuropsychological test data. Journal of Clinical and Consulting Psychology, 53 (6), 790-802.

American Psychological Association. (1984). Ethical Principles of psychologists and code of conduct. American Psychologist, 47, 1597-1611.

Barker, M. (1977). A short form of the Category Test. Psychological Reports, 40, 1243-1246.

Betram, K. W., Abeles, N., & Snyder, D. J. (1990). The role of learning in performance on Halstead's Category Test. The Clinical Neuropsychologist, 1 (3), 244-252.

Beaumont, J. G. (1975). The validity of the Category Test administered by on-line computer. Journal of Clinical Psychology, 31, 458-462.

Bertram, K. W., Abeles, N., & Snyder, P. J. (1990). The role of learning in performance on Halstead's Category Test. The Clinical Neuropsychologist, 4 (3), 244-252.

Bigler, E. D., Steinman, D. R., & Newton, J. S. (1982). Clinical assessment of cognitive deficit in neurologic disorder I: Effect of age and degenerative disease. Clinical Neuropsycholoy, 3 (3), 5-10.

Boll T. J. (1974). Behavioral correlates of cerebral damage in children aged 9 through 14. In R. M. Reitan and L. A. Davidson (Eds.), Clinical Neuropsychology: Current status and applications (pp. 601-675). Washington, DC: Hemisphere Publishing Corporation.

Boll T. J. (1981). The Halstead-Reitan Neuropsychological Battery. In S. B. Filskov & T. J. Boll (Eds.), Handbook of clinical neuropsychology. New York: Wiley-Interscience.

Boll T. J. (1993). Children's Category Test. San Antonio, TX: The Psychological Corporation.

Boll T. J. (1978). Diagnosing brain impairment. In B. B. Wolman (Ed.), Clinical diagnosis of mental disorders (pp. 601-675). New York: Plenum.

Boll, T. J., & Reitan, R. N. (1972). Comparative ability interrelationships in normal and brain damaged children. Journal of Clinical Psychology, 28, 152-156.

Boyle, G. L. (1975). Shortened Halstead Category Test. Australian Psychologist, 10, 81-85.

Boyle, G. L. (1986). Clinical neuropsychological assessment: Abbreviating the Halstead Category Test of brain dysfunction. Journal of Clinical Psychology, 42(4), 69-76.

Boyle, G. J. (1988). What does the neuropsychological Category Test measure? Archives of Clinical Neuropsychology, 3, 69-76.

Calsyn, D. A., O'Leary, M. R., & Chaney, E. F. (1980). Shortening the Category Test. Journal of Consulting and Clinical Psychology, 48, 788-789.

Cattell, R. B. (1963). Theory for fluid and crystallized intelligence: A critical experiment. Journal of Educational Psychology, 54, 1-22.

Chalfant, J. C., & Scheffelin, M. A. (1969). Central processing dysfunction in children: A review of the research. Bethesda: U.S. Department of Health, Education, and Welfare.

Charter, R. A., Adkins, T. G., Alekoumbides, A., & Seccat, G. F. (1987). Reliability of the WAIS, WMS, and Reitan Battery: Raw scores and standardized scores corrected for age and education. The International Journal of Clinical Neuropsychology, 9 (11), 28-32.

Choca, J. P., Laatsch, L., Wetzel, A., & Agresti. (1997). The Halstead Category Test: A fifty year perspective. Neuropsychology Review, 7 (2), 61-75.

Collins, M., & Odell, K. (1986). Computerization of a traditional test for nonverbal visual problem solving. Cognitive Rehabilitation, 4 (5), 16-18.

Committee to Develop Standards for Educational and Psychological Testing (1985). Standards for educational and psychological testing. Washington, DC: American Psychological Association.

CTB Macmillan/McGraw-Hill. (1988). California Achievement Tests, Forms E and F. Monetery, CA: Author.

Davidoff, G., Morris, J., Roth, E., Bleiberg, J. (1985). Cognitive dysfunction and mild closed head injury in traumatic spinal cord injury. Archives of Physical Medicine and Rehabilitation, 66, 489-491.

DeFilippis, N. A., & McCampbell, E. (1979). The Booklet Category Test: Research and clinical form. Odessa, FL: Psychological Assessment Resources, Inc.

DeFilippis, N. A., McCampbel, E., & Rogers, P. (1979). Development of a booklet form of the Category Test: Normative and validity data. Journal of Clinical Neuropsychology, 1, 339-342.

Dodrill, C. B., & Troupin, A. S. (1975). Effects of repeated administrations of a comprehensive neuropsychological battery among chronic epileptics. The Journal of Nervous and Mental Disease, 161 (3), 185-190.

Doehring, D. G., & Reitan, R. M. (1962). Concept attainment of human adults with lateralized cerebral lesions. Perceptual and Motor Skills, 14, 27-33.

Finlayson, M. A., Sullivan, J. F., & Alfano, D. P. (1986). Halstead's Category Test: Withstanding the test of time. Journal of Clinical and Experimental Neuropsychology, 8, (6), 706-709.

Fromm-Auch, D., Yeudall, L. T. (1983). Normative data for the Halstead-Reitan Neuropsychological Tests. Journal of Clinical Neuropsychology, 5 (3), 221-238.

Gardner, E. F., Madden, R., Rudman, H. C., Karlsen, B., Merwin, J. C., Callis, R., & Collins, C. S. (1989). Stanford Achievement Test, Seventh Edition Plus. San Antonio, TX: The Psychological Corporation.

Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). Measurement theory for the behavioral sciences. San Franscisco, CA: W. H. Freeman and Co.

Golden, C. J. (1978). Diagnosis and rehabilitation in clinical neuropsychology. Springfield: C.C. Thomas.

Golden, C. J., Kuperman, S. K., MacInnes, W. D. , & Moses, J. A. (1981). Cross-validation of an abbreviated form of the Halstead Category Test. Journal of Consulting and Clinical Psychology, 49, 606-607.

Golden, C. J., Hammeke, T.A., & Purisch, A. D. (1980). The Luria-Nebraska Neuropsychological Battery. Los Angeles: Western Psychological Services.

Golden, C. J., Osmon, D. C., Moses, J. A., Jr., & Berg, R. A. (1981). Interpretation of the Halstead-Reitan Neuropsychological Test Battery. New York, NY: Grune & Stratton.

Goldstein, G., & Shelly, C. H. (1972). Statistical and normative studies of the Halstead Neuropsychological Test Battery relevant to a neuropsychiatric setting. Perceptual and Motor Skills, 34, 603-620.

Gregory, R. J., Paul, J. J., & Morrison, M. W. (1979). A short form of the Category Test for adults. Journal of Clinical Psychology, 35, 795-798.

Halstead, W. C. (1940). Preliminary analysis of grouping behavior in patients with cerebral injury by the method of equivalent and non-equivalent stimuli. American Journal of Psychiatry, 96, 1263-1294.

Halstead, W. C. (1947). Brain and intelligence: A quantitative study of the frontal lobes. Chicago: The University of Chicago Press.

Halstead, W. C. , & Sattlage, P. H. (1943). Grouping behavior of normal persons with lesions of the brain. Archives of neurology and psychiatry, 49, 489-503.

Hartlage, L. C. (1981). Clinical application of neuropsychological data. School Psychology RevieSS, 10(3), 362-366.

Hartlage, & Long, C. J. (1997). Development of neuropsychology as a professional psychological speciality: History, training, and credentialing. In C. R. Reynolds, & E. Fletcher-Janzen (Eds.), The handbook of clinical child neuropsychology (pp. 3-16). New York: Plenum Press.

Hartlage, L. C., & Telzrow, C. F. (1983). Assessment of neuropsychological functioning. In Bracken & F. Paget (Eds.), Psychoeducational assessment of preschool and primary aged children (pp. 1-68). New York: Grune & Stratton.

Hays, W. L. (1981). Statistics (3rd ed.). New York: Holt, Rinchart, and Winston.

Hieronymus, A. N., Hoover, H. D., & Lindquist, E. F. (1986). Iowa Tests of Basic Skills, Forms G and H. Chicago: Riverside.

Horn, J. L. (1985). Handbook of Intelligence. New York: Wiley.

Horn, J. L. & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. Acta Psychologica, 26, 107-129.

Jastak, S. & Wilkinson, G. S. (1984). Wide Range Achievement Test-Revised. Wilington, DE: Jastak Associates.

Kilpatrick, D. G. (1970). The Halstead Category Test of brain dysfunction; Feasibility of a short form. Perceptual and Motor Skills, 30, 577-578.

Kimura, S. D. (1981). A card form of the Reitan-Modified Halstead Category Test. The Journal of Consulting and Clinical Psychology, 49, 145-146.

Kleinmutz, B. (1987). Automated interpretation of neuropsychological test data: Comment on Adams and Heaton. Journal of Consulting and Clinical Psychology, 55 (2), 266-267.

Klonoff, H., Robinson, C. G., & Thompson, G. (1969). Acute and chronic brain syndromes in children. Developmental Medicine and Child Neurology, 11, 198-213.

Klove, H. (1974). Validation studies in adult clinical neuropsychology. In R. M. Reitan and L.A. Davidson (Eds.), Clinical Neuropsychology; Current status and applications. Washington: Winston and Sons.

Laatsch, L. (1983). Development of a memory training program. Cognitive Rehabilitation, 1 (4), 15-18.

MacInnes, W.D., Forch, J.R., & Golden, C. J. (1981). Cross-validation of a booklet form of the Category Test. Clinical Neuropsychology, 3, 3-5.

Matarazzo, J. D., Matarazzo, R. G., Weins, A. N., Gallo, A. E., & Klonoff, H. (1976). Retest reliability of the Halstead Impairment Index in a normal, a schizophrenic, and two samples of organic patients. Journal of Clinical Psychology, 32 (2), 338-349.

Matarazzo, J. D., Weins, A. N., Matarazzo, R. G., & Goldstein, S. G. (1974). Psychometric and clinical test-retest reliability of the Halstead Impairment Index in a sample of healthy, young, normal men. The Journal of Nervous and Mental Disease, 158 (1), 37-49.

Matthews, C. G. (1974). Applications of neuropsychological test methods in mentally retarded subjects. In R. M. Reitan & L.A. Davidson (Eds.), Clinical neuropsychology: Current status and applications (pp. 267-287). New York: Hemisphere Publishing Corporation.

Mercer, W. N. (1993). Performance of brain-injured versus non-brain-injured individuals on three versions of the Category Test. Unpublished doctoral dissertation, University of North Texas, Denton.

Mercer, W. N., Harrell, E. H., Miller, D. C., & Rockers, D. (1997). Performance of Brain-Injured Versus Healthy Individuals on three versions of the Category Test. The Clinical Neuropsychologist, 2, 174-179.

Miller, D. C. (1989). Relationships among three levels of measuring planning: Electrophysiological - the Category Test, and psychological - PASS scales in adolescent males. Unpublished doctoral dissertation, Ohio State University, Columbus, OH.

Miller, D. C. (1994). The Computerized Version of the Category Test-Young Children's Version [Computer software]. Denton, TX.

Moerland, M. C., Aldenkamp, A. P., & Alpherts, W. C. J. (1986). A neuropsychological test battery for the Apple II-E. International Journal of Man-Machine Studies, 25, 453-467.

Moses, J. A. (1985). Internal consistency of standard and short forms of three itemized Halstead-Reitan neuropsychological battery tests. The International Journal of Child Neuropsychology, 7 (3), 164-166.

Morrison, P. R., & Gates, F. R. (1988). Assessment of a microcomputer-based version of the Austin Maze. The Journal of General Psychology, 115 (3), 307-314.

Nussbaum, N. L., & Bigler, E. D. (1989). Halstead-Reitan Neuropsychological Test Batteries For Children. In C.R. Reynolds & E. Fletcher-Janzen (Eds.), Handbook of Clinical Child Neuropsychology. New York, NY: Plenum Press.

Plaisted, J. R., Gustavson, J. W., Wilkening, J. G. N., & Golden, C. J. (1983). The Luria-Nebraska Neuropsychological Battery-Children's Revision. Theory and current research findings. Journal of Clinical Child Psychology, 12, 13-21.

Prescott, G. A., Barlow, I. H., Hogan, T. P., & Farr, R. C. (1985). Metropolitan Achievement Test, Sixth Edition. San Antonio, TX: The Psychological Corporation.

Raven, J. C. (1965). Guide to using the Colored Progressive Matrices. New York, NY: The Psychological Corporation.

Reed, H. B. C., Reitan, R. M., & Klove, H. (1965). Influence of cerebral lesions on psychological test performances of older children. Journal of Consulting Psychology, 29, 247-251.

Reeder, K. P., & Boll, T. J. (1992). A shortened intermediate version of the Halstead Category Test. Archives of Clinical Psychology, 7, 53-62.

Reitan R. M. (1955). An investigation of the validity of Halstead's measures of biological intelligence. AMA Archives of Neurological Psychiatry, 73, 28-35.

Reitan, R. M. (1969). Manual for administration of neuropsychological test batteries for adults and children. Indianapolis: privately published by the author.

Reitan R. M., & Davidson, L. (1974). Clinical neuropsychology: Current status and applications. Washington D.C.: Winston.

Reitan, R. M., & Wolfson, D. (1985). The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation. Tucson, AZ: Neuropsychology Press.

Reitan, R. M., & Wolfson, D. (1992). Neuropsychological evaluation of older children. South Tucson, AZ: Neuropsychology Press.

Reitan, R. M., & Wolfson, D. (1993). The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation (2nd ed.). South Tucson, AZ: Neuropsychology Press.

Reynolds, C. R., & Fletcher-Janzen, E. (1997). Handbook of Clinical Neuropsychology. New York, NY: Plenum Press.

Rockers, D. (1996). Test-Retest Correlation on the Computerized Version of the Category Test. Unpublished doctoral dissertation, University of North Texas, Denton.

Roth, E., Davidoff, G., Thomas, P., Doljanac, R., Dijkers, M., Berent, S., Morris, J., & Yarkony, G. (1989). A controlled study of neuropsychological deficits in acute spinal cord injury patients. Paraplegia, 27, 480-489.

Rothke, S. (1986). The role of set-shifting cues on the Wisconsin Card Sorting Test and the Halstead Category Test. The International Journal of Clinical Neuropsychology, 8 (1), 11-14.

Russell, E. W., & Levy, M. (1987). Revision of the Halstead Category Test. Journal of Consulting and Clinical Psychology, 55 (6), 898-901.

Sarazin, F. F. A., Spreen, O. (1986). Fifteen-year stability of some neuropsychological tests in learning disabled subjects with and without neurological impairment. Journal of Clinical and Experimental Neuropsychology, 8 (3), 190-200.

Sattler, J. M. (1992). Assessment of Children. San Diego: Jerome M. Sattler Publisher, Inc.

Seidenberg, M., O'Leary, D. S., Giordani, B., & Boll, T. J. (1981). Test-retest IQ changes of epilepsy patients: Assessing the influence of practice effects. Journal of Clinical Neuropsychology, 3, 237-255.

Sherrill, R. E. (1987). Options for shortening Halstead's Category Test for Adults. Archives of Clinical Neuropsychology, 55 (6), 898-901.

Shore, C., Shore, H., & Pihl, R. O. (1970). Correlations between performance on the Category Test and the Wechsler Adult Intelligence Scale. Perceptual and Motor Skills, 32, 70.

Spreen, O., & Strauss, E. (1991). A compendium of neuropsychological tests: Administration, norms, and commentary. New York, NY: Oxford University Press.

Summers, M. N., & Boyle, T. J. (1987). Comparability of a short booklet version and the traditional form of the Category Test. The International Journal of Clinical Neuropsychology, 9 (4), 158-161.

Taylor, D. J., Hunt, C., & Glaser, B. (1990). A cross validation of the Revised Category Test. Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2 (4), 486-488.

Taylor, J. M., Goldman, H., Leavitt, J., & Kleiman, K. M. (1984). Limitations of the brief form of the Halstead Category Test. Journal of Clinical Neuropsychology, 6 (3), 341-344.

Telzrow, C. F., & Hartlage, L. C. (1984). A neuropsychological model for vocational planning for learning disabled students. In W. Cruickshank & J. Kleibhan (Eds.), Early adolescents to early adulthood (pp. 143-156). Syracuse, New York: Syracuse University Press.

Tsushima, W. T., & Town, W. S. (1977). Neuropsychological abilities of young children with questionable brain disorders. Journal of Consulting and Clinical Psychology, 45, 757-762.

Vanderploeg, R. D., & Logan, S. G. (1989). Comparison of the Halstead Category Test and the Revised Category Test: Comment on Russell and Levy. Journal of Consulting and Clinical Psychology, 57 (2), 315-316.

Walsh, K. W. (1978). Neuropsychology: A clinical approach. Edinburgh: Churchill Livingstone.

Wechsler, D. (1974). Manual for the Wechsler Intelligence Scale for Children-Revised. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1991). Manual for the Wechsler Intelligence Scale for Children-Third Edition. San Antonio, TX: The Psychological Corporation.

Woodcock, R. W. (1977). Woodcock-Johnson Psycho-Educational Battery: Technical Report. Allen, TX: DLM Teaching Resources.

Yeudall, L. T., Reddon, J. R., Gill, D. M., & Stefanyk, W. O. (1987). Normative data for the Halstead-Reitan Neuropsychological Tests stratified by age and sex. Journal of Clinical Psychology, 43 (3) 346-367.

APPENDICES

TEXAS WOMAN'S UNIVERSITY
Debra J. Gomez - Principal Investigator
Daniel Miller, Ph.D. - Research Supervisor
(940) 898-2251

SUBJECT CONSENT TO PARTICIPATE IN RESEARCH

1. I hereby authorize Debra J. Gomez or one of the supervised graduate students to perform the following assessment procedures. My child will be administered a computerized version of the Category Test on two separate occasions approximately 3 weeks apart. Each administration will last approximately 15 minutes and my child will be asked to look at 80 images and individually categorize them. During the first session my child will also be administered the Concept Formation Test from the Woodcock-Johnson Tests of Cognitive Ability. This will last approximately 15 minutes and my child will be asked between 6-35 questions depending on ability.
2. These procedure have been explained to me in writing and reviewed orally by one of the researchers. The examiners will be one of three doctoral students who are under the supervision of Daniel Miller, Ph.D. My child will be administered the Computerized Version of the Category Test for Younger Children (CVCT-YC) which will take approximately 15 minutes and the Concept Formation test from the Woodcock-Johnson Tests of Cognitive Ability which will take approximately 15 more minutes. Exactly three weeks later the CVCT-YC will be re-administered. The CVCT-YC will require my child to look at a stimulus item presented on the computer screen and decide what color (red, blue, yellow, or green) best describes the picture. For example, a red circle is shown to my child and the correct answer would be red. If my child responds with the correct answer they receive feedback which tells them they are correct, and vise versa if an incorrect response is given. The 80 items are divided into five subtests. On the first four subtests there is one rule or principle which is the same across all items. The purpose of the Category Test is to measure my child's problem solving skills and cognitive flexibility. The last subtest is a memory test which will ask my child to recall correct answers from previous sections. My child will receive a certificate of participation for their efforts.
3. I understand that the procedures will primarily benefit the researchers who are attempting to validate a new test, the benefits to my child will be minimal.
4. The researchers will try to prevent any problem that could happen because of this research. I should let the researchers know at once if there is a problem and they will help me. I understand, however, that TWU does not provide medical services or financial assistance for injuries that might happen because my child is taking part in this research.
5. An offer to answer all of my questions regarding the study has been made. A description of the possible discomfort and risks reasonably expected have been discussed with me.

a. Confidentiality. Confidentiality of the research data will be maintained by assigning my child's record a case number. A master list which links the case number with my name and my child's name will be maintained in the Brain Behavior Laboratory office under lock and key.
b. School age children will be used as subjects. Consent will be obtained from both my child and myself. I understand that I or my child may terminate participation in the study at any time.
c. My child will be asked to participate in two fifteen minute testing sessions. Examiners have been trained to deal with possible test anxiety, fatigue, poor motivation, and other factors which may arise during testing.

6. The following list will be administered to my child:
   - The Computerized Version of the Category Test - Younger Children Version on two separate occasions.
   - The test of Concept Formation from the Woodcock-Johnson Tests of Cognitive Ability.

I agree to:

- Allow my child to be removed from the classroom (if testing is done during the school day)
- Allow my child to be individually administered, on two separate occasions, a computerized test and on the first occasion an orally delivered test. Both tests are designed to measure problem solving skills.

My child agrees to:

- Be individually administered a computer test (on two separate occasions) and an orally administered test on (on the first occasion) designed to measure problem solving skills.

Signature Required

__________________________ _______________

Parent/Guardian Date

Child's Birthdate (YY/MM/DD): __________ Today's Date: (YY/MM/DD): __________ Child's Sex: ______

Child's Grade in School: __________

Child's Ethnic Group (If the child has a blended ethnicity, use father's ethnic group):
____ African American
____ Asian American
____ Hispanic American
____ Native America
____ White Anglo American

Child lives with:
____ biological parents
____ adopted parents
____ step-parents
____ single parent
____ other (specify: ______________________)

Mean Family Income:
____ under $15,000 per year
____ $15,000 - $29,999 per year
____ $30,000 - $44,999 per year
____ $45,000 - $59,000 per year
____ $60,000 - $74,999 per year
____ $75,000 - $84,999 per year
____ $85,000 - $99,999 per year
____ over $100,000 per year

Marital Status of Parents or Guardians:
____ Married
____ Single Parent
____ Divorced
____ Widowed

Mother's Highest Educational Level:
____ less than High School
____ High School or GED
____ 2 yr Associate Degree
____ 4 yr Bachelors Degree
____ Master's Degree
____ Doctorate Degree

Father's Highest Educational Level;
____ less than High School
____ High School or GED
____ 2 yr Associate Degree
____ 4 yr Bachelors Degree
____ Master's Degree
____ Doctorate Degree