1997

A concurrent validity study of the Kaufman Brief Intelligence Test and the Wechsler Adult Intelligence Scale - Revised

Kimberly D. Rayburn

A CONCURRENT VALIDITY STUDY OF THE KAUFMAN BRIEF INTELLIGENCE TEST AND THE WECHSLER ADULT INTELLIGENCE SCALE-REVISED

BY

KIMBERLY D. RAYBURN

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS IN PSYCHOLOGY

MARSHALL UNIVERSITY GRADUATE COLLEGE

1997

Master of Arts Thesis

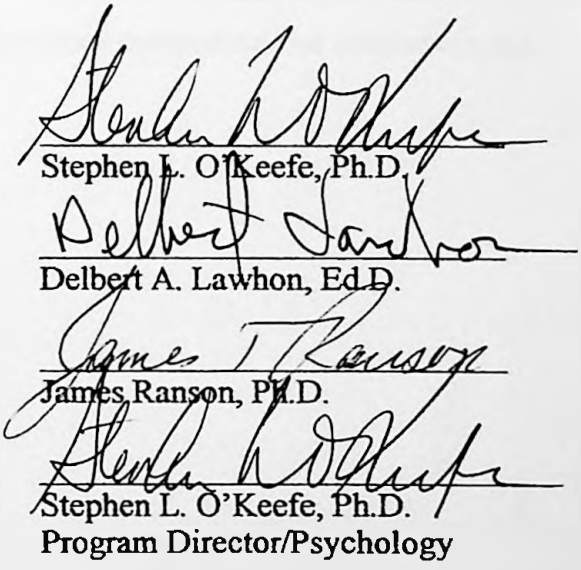
Of

Kimberly D. Rayburn

Approved:

Thesis Committee

Major Professor:



Stephen L. O'Keefe, Ph.D.

Delbert A. Lawhon, Ed.D.

James Ranson, Ph.D.

Stephen L. O'Keefe, Ph.D.
Program Director/Psychology

Marshall University Graduate College

1997

Abstract

The Kaufman Brief Intelligence Test (K-BIT) as a valid screening instrument for the more comprehensive Wechsler Adult Intelligence Scale-Revised (WAIS-R) was studied. Both assessment instruments were administered to 34 subjects referred to a private psychological practice as part of a requested battery of test. Each test produces a verbal score, nonverbal score and combined verbal and nonverbal score. Results indicate that each section of the K-BIT has a significant correlation to its WAIS-R counterpart. The correlation between the verbal scores, nonverbal scores and the combined scores of the two instruments were .87, .78, and .88, respectively. The K-BIT appears to be an adequate screening device for the WAIS-R when the more comprehensive measure is not necessary.

Acknowledgment

A very special thank you to Dr. Del Lawhon is extended for all his support and advice throughout comprehensive testing, practicum and finally the dreaded thesis. I would like to thank Dr. James Ranson for lifting the weight from my shoulders, by helping me with my statistics in such an uncomplicated manner. Many thanks to Dr. Stephen O'Keefe for forcing me to write a paper in APA style for Principles and Practice. It has helped me through the whole program and has made me look at writing in an entirely different light. I would also like to thank Amy Y. Simms for helping out with administration even when she had a really busy schedule.

This thesis would not have been possible without the never ending support of two people, Nina Shinaberry and Dr. David Clayman. Words seem so inadequate to express my thanks for the opportunity to pursue my dream. I offer both of you the utmost respect for your integrity and your commitment to the profession.

This thesis is dedicated to my family and friends for their patience and understanding when I failed to see the light at the end of a very long tunnel.

Table of Contents

List of Tables

The trend in psychological testing has varied throughout history. When Simon Binet and David Wechsler developed their comprehensive intelligence tests, the trend was for extensive testing. However, the utilization of psychological testing has declined but recently that is changing (Phares, 1992).

Utilization of Assessment Instruments

The resurgence of psychological testing is the result of the vast usefulness of assessment. Psychological assessment is used in a variety of settings. Academic testing is still used in the school setting to enable educators to make decisions regarding placement and academic planning. Psychological testing can also help rule out learning disabilities or attention problems.

Neurological testing is important to physicians to facilitate a diagnosis between organic and psychological problems. For example, the differential diagnosis between depression and dementia is supported by psychological testing. The utilization of psychological testing can also help estimate cognitive decline in clients that have degenerative diseases. Psychometric measures have been shown to determine soft tissue damage that may have been undetected by certain brain scans, because they appear to be more sensitive to changes. This type of information is very helpful when working with an individuals who have suffered traumatic brain injuries.

The use of forensic psychologists in the court systems has lead to a need for intellectual testing. Forensic psychologists are often asked to determine whether or not an individual is competent to stand trial or handle their own financial affairs, testing is frequently utilized to help answer these questions.

Limitations of Comprehensive Testing

Psychologists provide services to both industry and private clients. They are employed in a vast number of settings and are paid by a variety of sources. Just as the trend in testing has changed so has the reimbursement for services provided. The recent trend in mental health has been the development of managed care programs that handle insurance claims for the private practice setting. Managed care companies vary in what services they are willing to reimburse financially. The current trend has demonstrated a decline in reimbursements for assessment procedures. They fail to see the usefulness of testing for certain conditions.

The Wechsler Adult Intelligence Scale-Revised (WAIS-R) is currently the most popular and widely used assessment instrument (Phares, 1992). The WAIS-R is an example of a comprehensive intelligence test. Comprehensive measures of intelligence, even though very useful, do have several shortcomings. Instruments like the Wechsler scales require a lot of time to administer, score and interpret. They are given using a standardized format, which requires extensive training and expertise to handle non-standard situations. The test manual for administration reports 60 to 90 minutes as the average time range for the testing procedure alone. The test can be administered by a trained technician, but the scores must be interpreted by a trained psychologist. The strict qualifications of the test render it a very timely and costly instrument.

However, the decision to discount testing as useful has posed an ethical problem for psychologist that specialize in the area of assessment. Psychologists often perform testing to provide objective evidence to support a particular diagnosis. This objective support is often crucial for developing an accurate treatment plan. However,

psychologists who have chosen this area as a career, need to be paid for their time and expertise.

Screening Instruments

One viable alternative to comprehensive testing is the development of screening tests. A good screening device could be used in a more cost effective and timely manner. If a clinician had access to a reliable and valid screening instrument that had a good correlation with a more comprehensive instruments, then they could screen clients instead of using the comprehensive instrument. The screening instruments are not designed to replace comprehensive test, but they can be used to determine whether additional testing is necessary. Another use of an IQ screen would be to estimate an individual's IQ when only a range of functioning is required. The use of a screening instrument saves not only time for administration, but also money for supplies. A screening device could also help eliminate unnecessary testing.

There are several IQ screening devices available, but most of the devices are severely limited in their scope. The Slosson Intelligence Test (SIT) is one available instrument that is highly based on verbal abilities. The fact that it is verbally based limits its utilization. It excludes anyone that has a verbal difficulty and it is culturally biased, because of the emphasizes on language. The Raven's Progressive Matrix test is also a screening device. The matrix test is non-verbally based which renders its use with a greater variety of individuals and is less culturally biased. However, these two screens like many other screening instruments are either verbally or non-verbally based. A screening device with both a verbal and non-verbal section is more useful in general. For example, the Slosson Intelligence Test (SIT) may provide an estimate of functioning with

someone that experiences nonverbal difficulties and the Raven's Matrix Test may be useful with someone with verbal difficulties, but neither is useful for both situations.

Kaufman Brief Intelligence Test

Alan and Nadine Kaufman developed the Kaufman Brief Intelligence Test (K-BIT) with this concept in mind. The K-BIT contains two verbal sections and one non-verbal section. The verbal sections is comprised of an Expressive Vocabulary and Definitions sub-tests. The two scores together are used to generate a Vocabulary IQ. The non-verbal section of the test is entitled Matrices, which is also used to generate an IQ score. Administration of the K-BIT generates a Vocabulary IQ, Matrices IQ, and a Composite IQ. The manual for the K-BIT (1990) states that the test has the same mean and standard deviation as the WAIS-R. It also reports that the K-BIT Composite score and the WAIS-R Full Scale IQ have a correlation of ($r = .75$). However, the sample size only included 64 "normal" subjects between the ages of 16 to 47 years. Given the limited scope of the sample, additional research is required.

A subsequent study assessing the correlation between the WAIS-R and the K-BIT seems to indicate a high correlation between the two instruments. In a study conducted by Naugle, Chelune & Tucker (1993), they report the verbal, nonverbal, and combined measures of the instruments correlate at .83, .77, and .88, respectively. The sample consisted of 200 neurological patients. The subjects in the their study have been referred for broad range of neurological difficulties. In another study conducted by Parker (1993), he reports that typically the K-BIT Composite yields a score that is only four points less than the WAIS-R Full Scale IQ.

Given the fact that the "normal" subjects correlate with the WAIS-R at (.75) and the neurological patients yield a correlation of (.88) , the need for additional studies is crucial. Perhaps the difference between the subjects in the sample accounts for the variation in the magnitude of the correlation or it could be the result of the research design. Another possible explanation for the difference may be a result of where the damage lies. For example, global brain damage may present differently than localized damage. Additional research with a wider variety of subjects is necessary to confirm the results of previous studies conducted with a limited range of subjects.

The purpose of this study is to determine the concurrent validity of the K-BIT to the WAIS-R of referred patients in a private psychological practice. The subjects have been referred for a variety of reasons and include a large range of difficulties. This study will focus not only on the correlation between the K-BIT Composite Score and the WAIS-R Full Scale IQ, but will also compare the verbal and performance domains of both instruments. Given the results of the previous studies, a high correlation between the K-BIT and WAIS-R verbal domains and combined scores are expected, but there will likely be a difference for the performance domains.

Method

Subjects

Participants were 34 patients referred for psychological/neuropsychological assessment in a private psychological practice in southwest West Virginia. Subjects include 13 forensic evaluations, 13 custody evaluations, 3 neurological patients, 2 disability evaluations, 2 screens for learning disabilities and 1 cognitive screen to establish baseline functioning. The referral sources include court referrals, attorney

referrals, medical request and self-referrals. The subjects include 20 males and 14 females. The participants range in age from 17 to 71, with mean age of 35.8 years. Educational attainment varies among the subjects with 4 being the least amount of years obtained and 24 being the most years obtained with the mean years of education being 11.

Instruments

Kaufman Brief Intelligence Test (K-BIT). The K-BIT is individually administered to individuals ages 4 to 90 years. This test contains three sub-tests with include Expressive Vocabulary, Definitions and Matrices. The K-BIT contains a total of 130 test items and requires 15 to 30 minutes to administer. The testing materials include a manual, easel and individual test record. The nationwide standardization sample (N=2,022) was stratified on the variables of gender, geographic region, socioeconomic status, and race or ethnic group. The split-half reliability coefficient for the K-BIT Composite IQ ranges from (.88) to (.97) for the standardized sample by age (Kaufman & Kaufman, 1990).

The concurrent validity of the K-BIT has been reported to have a significant correlation with both the WAIS-R (r=.75) and with the Wechsler Intelligence Scale for Children-Revised (WISC-R) at (r=.80).

Procedures

The K-BIT and the WAIS-R were administered to subjects as part of a battery of test requested by the referring clinician to assess for the various referral questions posed. The tests were administered and scored by two psychometricians in the private practice. The tests were administered on the same day except on rare occasions when testing could not be completed during the same session. The tests were administered in the testing lab of

the private practice and also at the South Central Regional Jail when necessary. Two subjects had been given the WAIS-R previously, therefore, they were only administered the K-BIT.

Results

The Pearson Correlation Coefficient yielded a relatively high correlation for the individual domains of both instruments. The Vocabulary score of the K-BIT was found to correlate with the Verbal IQ of the WAIS-R at (r =.87). The Matrices index from the K-BIT correlates to its WAIS-R counterpart or Performance IQ at (r = .78). Results also indicate that the K-BIT Composite score and the WAIS-R Full Scale IQ correlate at (r =.88). The current results are somewhat higher than the correlations found between the domains in the standardization sample, which included 64 "normal" adults between the ages of 16 and 47. The correlation between the comparable indexes of the WAIS-R and K-BIT are shown in Table 1. There was essentially no differences noted in the obtained mean scores of the two instruments. The mean scores of the K-BIT were only slightly lower than the obtained means of the WAIS-R. The mean scores and standard deviations are shown in Table 2.

Discussion

The results of the study are nearly identical to the results obtained in the study by Naugle, Chelune and Tucker (1993), which included 200 neurological patients. Results of this study were as expected given the results of previous research. The verbal domains and the combined scores of the instruments correlate to a higher degree than the nonverbal domains of the two instruments.

Subjects were intellectually classified according to their Full Scale IQ from the WAIS-R and their obtained Composite Score from the K-BIT. Results indicate that the K-BIT properly classified 50 percent of the subjects. The K-BIT hit the correct classification in 17 out of 34 subjects, without the utilization of a confidence interval. The hits include one subject in the superior range of intellectual functioning, seven in the average range, five in the low average range, three in the borderline range and one in the mild mental retardation range. The misses include one subject in the very superior range of intellectual functioning, one in the average range, six in the low average range, seven in the borderline range and two in the mild mental retardation range as classified by the WAIS-R.

The hit rate improved by 26 percent, with the utilization of the 95th percent confidence interval, in conjunction with the Full Scale IQ of the WAIS-R. The use of a confidence interval increased the number of hits to 26 and decreased the number of misses to a total of eight. The utilization of the 95th confidence interval improved the hit rate in those subjects scoring in the low average and borderline ranges of intellectual functioning , as indicated by the WAIS-R.

The K-BIT has several features which make it a desirable instrument when compared to other screening measures, as well as, more comprehensive measures of intelligence. The K-BIT is relatively inexpensive, provides both a verbal and nonverbal measure of intelligence, and can be used with a large age range. It is also easy to administer and score, yet requires very little time.

One difference between the K-BIT and WAIS-R, is the fact that the definition section of the K-BIT Vocabulary domain requires the individual to not only read words but they

must also be able to spell certain words. The verbal sub-tests of the WAIS-R do not require the subject to read or spell. This difference in instruments may significantly impact the scores of individuals with reading problems. A limitation of this study is the fact that the subjects were not assessed for reading levels prior to examination.

An additional finding of the study is the intellectual functioning of the subjects. The mean scores obtained from the K-BIT and WAIS-R indicate that the majority of the subjects in the study were in low average range of cognitive functioning. The obtained means scores of the verbal, nonverbal and combined domains of both instruments fell nearly one standard deviation below the mean. The fact that the subjects in the study consistently performed in the low average range of cognitive functioning may limit the results of this study. However, it appears to adequately represent a typical clinical sample. Further research is necessary to determine whether or not the obtained results can be generalized to other populations. Perhaps, the K-BIT is an adequate screening device for individuals in the low average range of intellectual functioning but inadequate for individuals in the superior range of intellectual functioning.

References

Kaufman, A.S., & Kaufman, N.L. (1990). Kaufman Brief Intelligence Test . Circle Pines, MN: American Guidance Service.

Naugle, R.I., Chelune, G.J., & Tucker, G.D. (1993). Validity of the Kaufman Brief Intelligence Test. Psychological Assessment, 5(2), 182-186.

Parker, L.D. (1993). The Kaufman Brief Intelligence Test: an introduction and review. Measurement and Evaluation In Counseling and Development, 26, 152-155.

Phares, E.J. (1992). Clinical psychology concepts, methods and profession (4th. ed.). Belmont, CA: Brooks & Cole Publishing Company.

Table 1. Pearson Correlation Between the Comparable Indexes of the WAIS-R and K-BIT

		Index		
Measures	N	Verbal	Nonverbal	Overall
WAIS-R /K-BIT	34	.87	.78	.88

Note. All correlations are significant at *p < .0001. WAIS-R = Wechsler Adult Intelligence Test-Revised ; K-BIT = Kaufman Brief Intelligence Test.

Table 2. Mean Scores and Standard Deviations of the WAIS-R and K-BIT

Scales	N	Mean	SD
Verbal IQ	34	86.76	16.81
Vocabulary	34	86.53	17.43
Performance IQ	34	88.32	16.80
Matrices	34	84.50	19.20
Full Scale IQ	34	86.73	16.97
Composite	34	84.06	19.37

Note. *p< .0001

Appendix A: Review of Literature

The ability to measure an individual's intellectual capacity has been the goal of numerous psychologists throughout history. The intangible concept of intelligence has been defined and assessed in a variety of ways, depending upon the beliefs of the clinician. Throughout the development of instruments to assess intelligence, clinicians have been asked to estimate an individuals intellectual capacity for a multitude of reasons. The way in which the psychologist attempted this task relied upon their theory of what comprises intelligence. The question of what constitutes intelligence has generated countless theories and assessment devices.

History of Assessment

Several theories of intelligence have been proposed to provided an operational definition of the concept. The theories range from the origin of intelligence to the definition of intelligence. Another obstacle in the path of assessment instruments has been the variations in cultures. Is there a way to measure intellectual functioning that is cultural-free? Psychologists have tried to bridge the language barrier as well as other culture differences that exist such as history.

Theorists have debated for years over the origin of intelligence. Some clinicians believe that intelligence is genetic and thereby inherited from an individual's parents. While others believe that the intellectual capacities of an individual is directly related to the stimulation in their environment. However, neither theory is universal in practice because there are exceptions to both theories. For example, there are children that develop to a higher level of cognitive functioning than their parents, with the reverse also being true in some cases.

Research with twins seems to indicate a strong genetic influence. Studies conducted with monozygotic, dizygotic, biological siblings, and adopted children suggest that both environment and genetics play a role in intellectual functioning. Genetic factors account for approximately 50 percent of an individual's capacities.

Studies also suggest that vocabulary is the best indicator of intelligence. The number of words that an individual learns is correlated to their ability to learn information. Therefore, it is a good measure of intelligence (Sattler, 1992). However, if an individual's vocabulary is the best estimate of their intelligence, then how can any one test to measure intellectual ability be cultural-free?

Theories of Intelligence

The first theory of intelligence was developed by Sir Francis Galton and J. McKeen Cattell. They developed a theory based on the keen sensory abilities. The theory proposed that everything an individual learns is related to their senses because all information is gained through the senses. The theory was discounted quickly, but it has continued to influence theorist throughout history (Gregory, 1996).

Another theory of intelligence was proposed by Charles Spearman. He suggested that two types of intelligence exist, a general intelligence or "g" and a specific intelligence. According to Spearman, "g" or general ability is related to such abilities as deductive reasoning, speed and the limit of an individual's intellectual functioning (Sattler, 1992). He believes that an individual's functioning consists of their general ability which "involves mainly the education of relations and correlates" (Gregory, 1996). The general ability, with one or two other specific traits, constitutes their intelligence (Sattler, 1992).

T.G. Thurstone also developed a theory of intelligence. He believed that individuals possess primary mental abilities. These abilities consist of verbal comprehension, word fluency, number, space, associative memory, perceptual speed, and inductive reasoning (Gregory, 1996). Thurstone disagreed with the idea that a single general factor accounted for abilities. He believed that a group of factors not a single general ability explained functioning (Gregory, 1996).

Raymond Cattell and John Horn generated a theory of intelligence based on the idea that there is a crystallized and fluid intelligence. They were of the school of thought that both environment and genetics contribute to an individual's ability. They credit an individual's fluid intelligence to inherited factors. The concept of fluid intelligence describes those skills that are nonverbal and cultural-free (Sattler, 1992). Crystallized intelligence is very dependent on culture, because it is considered to contain information that is learned. This type of intelligence is gained by applying fluid intelligence in a particular setting, such as school (Gregory, 1996).

Another theorist was H. Gardner. He developed the theory that several independent intelligence's exist. He defined intelligence as "the ability or skill to solve problems or to fashion products which are valued within one or more cultural settings" (Gregory, 1996). He studied brain-behavior relationships and proposed the idea that several different types of intelligence such as linguistic, musical, logical-mathematical, spatial, bodily-kinesthetic and personal exist (Gregory, 1996).

Assessment Instruments

There have been several theories of intelligence proposed throughout history, but only a few theorists have been successful at measuring "intelligence." The two most widely

accepted instruments were developed by Alfred Binet and David Wechsler. Both theorists developed comprehensive instruments of assessment that are still used today.

Defining the concept of intelligence is a difficult task, but developing a method to measure the abstract construct is even more difficult. Alfred Binet and Theodore Simon are credited with developing the first widely accepted intelligence test in 1905. Binet and Simon defined intelligence as the "ability to judge well, to understand well, to reason well" (Gregory, 1996). The test that they developed was entitled the Binet-Simon and was based on the assumption that intelligence is genetic, but that it can be affected by environmental factors.

The development of the Binet-Simon was in response to the changes in the progressing educational system in both Europe and North American (Phares, 1992). Since the first edition of the instrument, it has been revised four times. The first four editions of the Binet-Simon were similar in content with the fifth edition being quite different that its predecessors.

The first four versions of the Binet-Simon were created with the use of age scales. There were 20 age levels and each level contained six test items (Phares, 1992). The items passed were converted into a certain number of mental months depending upon the level. This number was then used to compute the individual's mental age or IQ.

In 1986, the test underwent a number of changes. The Stanford-Binet fifth edition contained four main categories of items. Each category of the test was designed to measure a different type of ability. The main abilities included verbal reasoning, quantitative reasoning, abstract/visual reasoning and short-term memory (Phares, 1992). The categories are divided into a number of sub-tests created to measure each particular

ability area. According to Phares (1992), the test uses a multistage approach to testing. This approach allows the practitioner to skip easier items when working with a bright individual.

Administration of the Standford-Binet generates standard scores in the four main ability areas along with a composite score. These scores are used to compute the individual's mental age or IQ.

Development Of the Wechsler Scales

Critics of the Binet-Simon cited a number of disadvantages of the test that led to the develop of the Wechsler scales. The Binet-Simon was criticized because it was primarily developed for use in the school setting and that it is too focused on verbal skills. David Wechsler was interested in developing an instrument to measure intelligence on a wider variety of people and abilities. In addition to an individual's verbal skills, he was also interested in non-verbal skills. In spite of the criticism, Wechsler was influenced by the work of Binet. He was also influenced by the work of Robert Yerkes. Yerkes is best known for the development of the Army Alpha and Beta tests used by the military for selection purposes during World War I (Gregory, 1996).

David Wechsler was interested in developing an instrument that could be used to test the psychiatric patients in Bellevue Hospital in New York (Phares, 1992 ; Gregory, 1996). Wechsler viewed intelligence as "the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with the environment" (Gregory, 1996). In 1939, he developed the Wechsler-Bellevue. The test contained two separate scales, a Verbal scale and a Performance scale. Each scale is comprised of separate sub-tests. The Verbal scale contained six sub-tests and the

Performance scale contained five. The scales both generate an IQ score and the two scales together generate a Full-Scale IQ.

The Wechsler-Bellevue was developed for use with adults. Wechsler's goal was to make it more interesting and motivating to adults than the school based Binet-Simon (Phares, 1992). The instrument was developed with eleven different sub-tests, with each sub-test arranged from the most simple items to increasingly more difficult items. One of the most useful characteristic of the test was the inclusion of the Performance scales. This addition made it easier to test individuals with verbal problems. The Performance IQ is also a better estimate of intelligence when there is a cultural difference, due to the fact that it does not require the utilization of language.

Wechsler believed that intelligence was evenly distributed among the population and he assessed this by using a "deviation IQ". He believed that one should interpret an individual's IQ by comparing them to their peers. Phares (1992) states that "this method statistically establishes an IQ of 100 as the mean for each age group." This concept remains the same for each norm group regardless of age.

Administration of Wechsler-Bellevue generated a scaled score for each one of the eleven sub-tests, a Verbal IQ score, Performance IQ score and a Full-Scale IQ score. Statistically the mean for the IQ scores is 100 with a standard deviation of 15. The scaled scores have a mean of 10 with a standard deviation of 3.

A clinician can estimate an individual's IQ by comparing their obtained scores to those of their peers. The range of intellectual functioning is determined by establishing the deviation from the mean. For example, individuals in the very superior range of

intellectual functioning obtain scores that are two standard deviations above the mean or a score of 130.

Since the Wechsler-Bellevue was developed, the test has been revised two times. It was revised in 1955 and again in 1981, with very few changes to the original test. A few items were added or deleted to make the test more culturally fair, but the test as a whole has remained the same. The name of the test was changed to the Wechsler Adult Intelligence Scale. The third revision of the test is being released in the summer of 1997. According to Phares (1992), both the adult and children's versions of the Wechsler scales "have become the most widely used techniques to assess intellectual functioning."

Utilization of Screening Instruments

Given the limitations of comprehensive intelligence tests, the K-BIT as an independent screening instrument has been the focus of research. According to Naugle, Chelune & Tucker (1993), the K-BIT appears to be an acceptable screening measure for verbal, nonverbal and combined intellectual abilities. This finding is in agreement with a similar study conducted by Parker (1993). Both studies indicate that the K-BIT is a useful instrument within certain boundaries. The use of the K-BIT is appropriate in certain situations. A situation which warrants the use of the K-BIT is "when time constraints preclude use of a longer measure" (Naugle et al.,1993). Parker (1993) feels that the test is useful in more than one way. He reports that the test is appropriate in a number of different situations. For example, screening several children for learning difficulties in a short period of time, estimating intelligence within a battery of personality instruments, determining educational diagnosis and especially as a follow-up to previous testing.

Naugle, Chelune & Tucker (1993) hoped to replicate the original correlation between the K-BIT and WAIS-R with another normative group. According to Kaufman & Kaufman, (1990) the original sample included "normal" subjects, which may have compromised the usefulness with other normative groups. Naugle et al.(1993) conducted a study of 200 neuropsychological patients to determine the correlation between the K-BIT and WAIS-R.

The results of the study indicate a high correlation between the two instruments across all sub-tests within their given domain. The verbal sub-tests of the K-BIT and WAIS-R correlated at (r = .83). The nonverbal sub-tests of the two instruments correlated at (r = .77). The Composite IQ of the K-BIT and Full Scale IQ of the WAIS-R correlated at a (r = .88) overall. The difference between the "nonverbal scores of the two measures, although statistically significant, are more varied across the age groups"(Naugle, Chelune & Tucker, 1993). One possible explanation for this difference may be the fact that the nonverbal performance on the WAIS-R is considerably effected by processing speed and visual-motor coordination. Further research on this difference may provide additional support for the research conducted by Wang & Kaufman (1993) in regards to intellectual aging patterns, which will be reviewed in more detail later . Another factor to consider is the fact that the K-BIT "relies less heavily on verbal output than does the WAIS-R" (Naugle et al., 1993).

Comparison of the Verbal and Nonverbal Dichotomy

An additional finding of the study (Naugle, Chelune & Tucker, 1993) was that the K-BIT yielded comparable scores to the WAIS-R counterparts but obtained score on the K-BIT were consistently higher. The K-BIT yielded verbal IQ scores that were 3.11 points

higher than the WAIS-R. The nonverbal scores were typically 5.19 points higher. The Composite IQ and Full Scale IQ differed by an average of 4.27 points. The research indicates that even though the average difference between the Composite IQ and Full Scale IQ are similar there were several variations among the sample. For example, the K-BIT Composite IQ varied from 12 points less than to 22 points higher than the Full Scale IQ of the WAIS-R.

Parker's (1993) analysis of the normative data contradicts the findings of Naugle, Chelune & Tucker (1993) in regards to obtained score differences. The Composite IQ of the K-BIT was on the average 4 points less than the WAIS-R Full Scale IQ and 6 points less than the WISC-R Full Scale IQ.

Results obtained by Parker (1993) are consistent with Naugle, Chelune & Tucker (1993) in regards to the Verbal and Performance dichotomy. Even though the obtained scores on the K-BIT correlate highly with obtained scores on the WAIS-R other factors may need to be considered. Parker (1993) reports that the sub-tests selected for the K-BIT were intended to correspond to the Wechsler's Verbal and Performance dichotomy. The K-Bit measures verbal abilities through expressive vocabulary and definitions, Parker (1993) questions whether or not this is enough to accurately measure verbal comprehension skills.

The Verbal and Performance dichotomy was further studied by Burton, Naugle & Schuster (1995). The research indicates that Parker (1993) may have had valid concerns regarding the relationship between the verbal and nonverbal domains. The study conducted by Burton et al. (1995) suggests that the "K-BIT Verbal Intelligence factor appeared to have a significant visuospatial component, calling into question the

assumption that the WAIS-R and K-BIT provide equivalent measures of verbal intelligence". Their study included 198 neurological patients with a vast array of difficulties including seizure disorders, closed head injuries, substance abuse, psychiatric disturbance, cerebrovascular disease, dementia, HIV dementia, intracerebral tumor, multiple sclerosis, toxic exposure and hydrocephalus.

The study designed several different model in which to study the relationship between the two instruments. Results seems to indicate that there is a difference in what the two instruments measure. Burton, Naugle & Schuster (1995) report that "the measure of verbal intelligence provided by the K-BIT does not appear to be empirically equivalent to the measure of verbal intelligence provided by the WAIS-R". The findings also indicate that the K-BIT "provides less of a differentiation between the verbal and nonverbal intellectual functions compared to the WAIS-R" (Burton et al., 1995).

Burton, Naugle & Schuster (1995) also report that given the visuospatial component of the Vocabulary sub-test of the K-BIT the Verbal IQ may provide scores that are spuriously low estimates of verbal intelligence. This problem may conceal discrepancies between K-BIT Vocabulary and Matrices IQ's.

Overall the results of the study indicates that the K-BIT is "well designed, adequately normed, and clinically useful screening battery for assessing intelligence when it is not practical to administer the full WAIS-R" (Burton, Naugle & Schuster, 1995).

Intelligence Research

The research on intelligence includes a wide variety of topics and interests. One study conducted evaluated subject differences such as race, gender, age and educational attainment as measured by the K-BIT. Another area of interest to researchers has been

the changes in IQ across the life span. Several studies have been conducted to determine the appropriateness of the K-BIT as a screening device for the Wechsler scales for children, as well as, adults. The WISC-R and WISC-III, were studied with different populations of children. The studies include children referred for poor academic progress, gifted and juvenile offenders.

Effects of Gender, Race, Age and Educational Attainment

The goal of a good intellectual assessment instrument is to eliminate as much bias as possible. In a study conducted by Kaufman & Wang (1992), they explored group differences in regards to gender, race and educational attainment. The main purpose of the study was to examine crystallized and fluid abilities across the variables of gender, race and socioeconomic status as determined by educational attainment. Another objective of the study was to explore the changes in abilities across the life span. The crystallized abilities were measured based on the individuals obtained score on the Vocabulary sub-test, which estimates acquired knowledge and verbal abilities. The fluid abilities were assessed by examining the individuals obtained scores on the Matrices sub-test, which attempts to measures simultaneous processing and nonverbal abilities.

The normative sample of the K-BIT assessment instrument was employed to provide the necessary data. For subjects between the ages of 4 and 19, the parent's educational attainment was utilized to determine years of education on behalf of the subject.

The results of the study indicate that gender differences were minimal. In regards to educational attainment, the highest educated subjects scored better than the lowest educated subjects. Racial differences were also evident. The White subjects outscored the Black subjects by nearly a whole standard deviation. Racial differences also included

variations between Whites and Hispanics. The Whites tended to outscore the Hispanics by a greater margin on the verbal sub-tests than the nonverbal sub-test. The Matrices sub-test produced no significant differences between the two groups across all age levels. The Hispanic group typically outscored the Blacks at younger age levels, but not at the adolescent and adult levels.

Educational attainment at all age levels significantly effected the obtained scores. Differences were noted on both the Vocabulary and Matrices sub-tests, with the Vocabulary sub-tests being the most effected by education. Subjects with more education consistently outscored the less educated subjects.

Intellectual Aging Patterns

The normative sample was utilized in yet another study by Wang & Kaufman (1993), in which they assessed changes in fluid and crystallized intelligence across the ages of 20 to 90. The study included 500 adult subject from the 2,022 subjects employed in the original sample. The purpose of the study was to examine the intellectual aging pattern across the life span. Wang & Kaufman (1993) suggest that the K-BIT may provide a more accurate estimate of this construct than the WAIS-R. This conclusion is based on the idea that the nonverbal sub-test of the WAIS-R requires visual-motor coordination and processing speed where the K-BIT relies on neither factor. This theory renders the K-BIT a more useful instrument in assessing fluid reasoning.

In regards to the intellectual aging pattern the belief is that individuals typically retain their crystallized abilities to a greater degree than their fluid abilities. This study yielded similar results to studies previously conducted with the WAIS and WAIS-R. The current study indicates that an individuals crystallized abilities typically plateau in their

40s and begin to decline in their late 60s and early 70s. The fluid abilities typically follow a different aging pattern. Research indicates that an individual's fluid intelligence peaks during late adolescents and then declines across the life span (Wang & Kaufman, 1993). This phenomena appears to be consistent regardless of educational attainment.

The Wechsler Scales for Children

One advantage of the K-BIT is the fact that it may be used with individuals between the ages of 4 to 90 years. The WAIS-R is only appropriate with individuals 16 to 74 years of age. This restriction requires the use of another test with individuals younger than 16 years of age. Wechsler designed tests for both adults and children. The appropriateness of the K-BIT with children has also been an area of interest to researchers.

Children with Academic Problems

In a study by Prewett (1992b) 35 subjects referred for poor academic progress were tested with both the WISC-R and the K-BIT to determine usefulness of the K-BIT as a screening device when the WISC-R is the comprehensive measure of intelligence utilized. The average subject in the sample obtained a Full Scale IQ that was more than one standard deviation below the mean, indicating that the majority of the subjects were functioning below average.

The results indicate that the two instruments share 66% variance, which implies that the two test measure comparable constructs. Prewett (1992b) reports that the verbal domains of the K-BIT and WISC-R correlate at a ($r = .83$), the performance domains at ($r = .70$) and that the Composite Score and Full Scale IQ correlate at ($r = .81$). The results also indicate that even though the two instruments correlate to a high degree, the

K-BIT Composite is frequently less than the WISC-R Full Scale IQ. Prewett (1992b) hypothesizes that the discrepancy may be explained, in part, by the variance in the performance domains. The research indicates that the K-BIT typically yields scores that are 6.2 points less than the WISC-R. Prewett (1992b) reviewed the possibility that due to this discrepancy the K-BIT may overidentify students in need of further assessment. He also hypothesized that "the relationship between the K-BIT and WISC-R may vary across populations"(Prewett, 1992b).

In another study conducted by Prewett (1995) he compares the WISC-III to both the K-BIT and the Matrix Analogies Test-Short Form (MAT). The subjects included 50 students referred for evaluation due to unsatisfactory academic progress. Prewett (1995) found that both the K-BIT and MAT correlated to a high degree with the WISC-III, (.78) and (.67) , respectively. He found that both screening instruments yielded somewhat higher scores than the more comprehensive measure. The K-BIT was found to be typically 4.8 points higher than the WISC-III. In the previous study conducted by Prewett (1992b), the more comprehensive measure typically yielded the highest scores. Prewett (1995) recommends that the clinician be familiar with both the screening and comprehensive measure in regards to which instrument yields the highest scores. Given the difference between the two studies, this information is crucial when trying to determine whether additional testing is necessary. Familiarity with the test could help decrease the possibility of overidentifying or underidentifying students at risk. Based on the result of the study Prewett (1995) recommends both the K-BIT and the MAT as valid screening instruments when the WISC-III is the criterion measure. As with the previous

study (Prewett, 1992b), Prewett cautions that the size and demographic characteristics of the sample may restrict the generalization to other populations.

Gifted Evaluations

Levinson & Folino (1994) conducted a study with children referred for gifted. The purpose of the study was to compare the K-BIT to the WISC-III with a gifted population. The obtained results were very different than the findings with students experiencing academic difficulties (Prewett,1995). The study (Levinson & Folino, 1994) found the correlation between the two assessment instruments to be less. The Verbal, Performance and Full Scale IQ correlate at .35, .35, and .53, respectively. These results give support to Prewett's (1995) concern regarding generalizations to other populations. Levinson & Folino (1994) convey that the "small sample size and lack of a counterbalanced administration of tests may have affected the results". Another explanation offered for the obtained difference is the fact that the WISC-III yields somewhat lower scores than the WISC-R. This information would, in part, explain the difference between this study and the Prewett (1992b) study.

Juvenile Delinquent

Prewett (1992a) conducted another study involving the relationship between the K-BIT and WISC-R. The sample included 40 male juvenile defenders that were academically deficient. The subjects in this study were also in the low average range of intellectual functioning. Only a small mean score difference was obtained between the two test. The difference was a nonsignificant .45 points. The results indicate that the K-BIT Composite score and the WISC-R Full Scale IQ correlate at ($r = .64$). This finding is

different than Prewett's (1992b) other study assessing children with poor academic progress, in which the two instruments correlated at (r = .81).

The observed difference between these three studies (Prewett, 1992a, 1992b, ; Levison & Folino, 1994; in press-a, in press-b) provides further support for the concern that large discrepancy exist between different populations. All three samples yielded different results suggesting that the clinician must select the screening instrument based on the given population.

Bibliography

Burton, D. B., Naugle, R.I., & Schuster, J.M. (1995). A structural equation analysis of the Kaufman Brief Intelligence Test and the Wechsler Adult Intelligence Scale-Revised. Psychological Assessment, 7(4), 538-540.

Gregory, R.J. (1996). Psychological testing history, principles, and applications (2nd. ed.). Needham Heights, MA: Allyn & Bacon.

Kaufman, A.S., & Kaufman, N.L. (1990). Kaufman Brief Intelligence Test. Circle Pines, MN: American Guidance Service.

Kaufman, A.S., & Wang, J.J. (1992). Gender, race, and education differences on the K-BIT at ages 4 to 90 years. Journal of Psychological Assessment, 10, 219-229.

Levinson, E.M., & Folino, L. (1994). The relationship between the WISC-III and the Kaufman Brief Intelligence Test with students referred for gifted evaluation. Special Services in the Schools, 8(2), 155-159.

Naugle, R.I., Chelune, G.J., & Tucker, G.D. (1993). Validity of the Kaufman Brief Intelligence Test. Psychological Assessment, 5(2), 182-186.

Parker, L.D. (1993). The Kaufman Brief Intelligence Test: an introduction and review. Measurement and Evaluation in Counseling and Development, 26, 152-155.

Phares, E.J. (1992). Clinical psychology concepts, methods, and profession (4th. ed.). Belmont, CA: Brooks & Cole Publishing Company.

Prewett, P.N. (1992a). The relationship between the Kaufman Brief Intelligence Test (K-BIT) and the WISC-R with incarcerated juvenile delinquents. Educational and Psychological Measurement, 52, 977-981.

Prewett, P.N. (1992b). The relationship between the Kaufman Brief Intelligence Test (K-BIT) and the WISC-R with referred students. Psychology in the Schools, 29, 25-27.

Prewett, P.N. (1995). A comparison of two screening tests (the Matrix Analogies Test-Short Form and the Kaufman Brief Intelligence Test) with the WISC-III. Psychological Assessment, 7(1), 69-72.

Sattler, J.M. (1992). Assessment of children (Rev. 3rd. ed.). San Diego, CA: Jerome M. Sattler, Publisher, Inc.

Wang, J.J., & Kaufman, A.S. (1993). Changes in fluid and crystallized intelligence across the 20- to 90-year age range on the K-BIT. Journal of Psychoeducational Assessment, 11, 29-37.