CLINICAL AND FORENSIC ISSUES REGARDING AGE, EDUCATION, AND THE VALIDITY OF NEUROPSYCHOLOGICAL TEST RESULTS: A REVIEW AND PRESENTATION OF A NEW STUDY

Ralph M. Reitan
Deborah Wolfson

Send a comment to the authors about this paper

Accepted for publication in the *Journal of Forensic Neuropsychology*

Abstract

This paper reviews procedures and reports of adjustments of raw scores from the Halstead-Reitan Neuropsychological Test Battery for Adults using the method proposed by Heaton, Grant, and Matthews (HGM) (1991). Arguments and findings both supporting and criticizing the HGM transformations were reviewed.

From a psychometric viewpoint, the conclusion was reached that the HGM method, which is based only on a neurologically-normal group, transforms raw scores into scaled scores that tend to fall in the normal range, especially for brain-damaged persons. Although this question is not yet adequately researched, the effect may stem principally from adjustments made for persons with less education and older age. The consequent result of the adjustments, in this case, would be to produce normal scores for tests that were designed, developed, and validated as measures of brain impairment - defeating the very purpose of the tests as neuropsychological measures. If the transformation process, based on age and education, is faulty for any reason (including inadequacies of the normative sample), and the transformations skew the raw scores either in the direction of normalcy or brain impairment, a serious problem exists in clinical interpretation or in a forensic setting. Thus, the questions raised in this paper are of theoretical, clinical, and forensic importance, and are urgently in need of resolution through further empirical research.

Heaton, Grant, and Matthews presumed, quite reasonably, that persons with more education would perform better on neuropsychological tests than would persons with lesser education, and that, among adults, older people might show some deterioration of performances in comparisons with younger subjects. A considerable body of published evidence had already supported these postulates. Thus, it was appropriate to explore the prospect of developing adjustments of raw scores in accordance with the age and education of the individual subject to evaluate the subject's performances with relation to the rest of the population.

Although relatively little data were available to determine whether these postulates were also true for brain-damaged subjects, this was not an immediate problem, inasmuch as Heaton, Grant and Matthews intended to develop the age and education adjustments using a sample of neurologically-normal persons. Their research supported the hypotheses underlying the purposes of their investigation, and was published in a volume that presented extensive tables for transforming raw scores on a series of tests to T-scores having a mean of 50 and a standard deviation of 10 for each test (Heaton, Grant, & Matthews, 1991). Jarvis and Barth (1994), in their book that was largely devoted to instruction in clinical interpretation of Halstead-Reitan test results, stated that the Heaton, Grant, and Matthews T-score transformations were generally used in the illustrative cases that they presented.

There can be no doubt that these transformations are used widely, and presently often serve as a basis for clinical and forensic interpretations of results for individual tests, for patterns and relationships among test results, and for conclusions about individual patients and clients, even though there has been very little research comparing the clinical significance of the transformations with clinical interpretations based on the raw scores. Obviously, the aim and purpose of the profession is to produce valid conclusions about the effects of brain damage or impairment for the individual examinee. Our integrity as neuropsychologists rests upon the accuracy with which we achieve this aim. From a forensic viewpoint, the validity of settlements and judgments may be determined by the accuracy of neuropsychological conclusions. Whenever raw test scores are adjusted or transformed, it is obvious that the profession must have full and complete information about the effects of such changes.

Many clinical neuropsychologists were well aware that neuropsychological tests, because of their specified purpose, have been developed to meet conditions of sensitivity and specificity - sensitivity to brain damage or impairment and specificity for brain impairment. If a test is useful as a neuropsychological instrument, the test scores should reflect brain impairment and should not overlap significantly with other groups, including non-brain-damaged persons.

What are the practical implications of these conditions? When they are met, brain impairment becomes a determiner of the test results, at least to a large extent, among persons with brain damage, and the remaining variance, determined by attribute variables such as age and education, is correspondingly reduced. It is only reasonable to postulate that age and education, which vary relatively randomly among brain-damaged persons, will lose a major aspect of their significance if brain impairment is a major factor in producing scores on neuropsychological tests (as should be the case if the tests are, in fact, *neuro*psychological).

While adjustments for age and education might be justified among non-brain-damaged persons, such adjustments among brain-damaged persons should have lessened justification. In fact, adjustments of raw scores of brain-damaged persons, when based on data from non-brain-damaged subjects, might possibly interfere with the validity of raw scores as indicators of brain impairment. Conceivably, such adjustments, based on trends and relationships that exist in non-brain-damaged groups, might actually have their major effect by influencing scores of brain-damaged persons to move in the direction of the scores of non-brain-damaged persons, inasmuch as these trends (correlations) might not exist in the brain-damaged group.

Considerations of this kind led Reitan and Wolfson to conduct a series of studies investigating the correlation between age and education and neuropsychological data using groups without brain damage and groups with brain damage (as well as other conditions which might reflect brain impairment, such as learning disabilities). The General Neuropsychological Deficit Scale (GNDS) was selected as the variable to reflect the adequacy of brain functions since the GNDS score, based on 42 brain-sensitive variables summarizing the results of the Halstead-Reitan Battery, has been found to be sensitive to impairment of brain functions (Reitan & Wolfson, 1988, 1995a), and even more sensitive than the Impairment Index (Rojas & Bennett, 1995; Oestreicher & O'Donnell, 1995).

The first study (Reitan & Wolfson, 1995) used two approaches to examine the effects of age and education on GNDS scores: (1) Were there significant differences in level of performance between older and younger subjects, and (2) Were coefficients of correlation between age and education versus GNDS scores statistically significant? The research answered these questions positively in both instances for the non-brain-damaged group, but negatively for the group with definite evidence of brain damage or disease. The effects of brain damage had subverted the significant influence of age and education that was apparent in the control group.

The findings definitely implied that brain-damaged and control groups cannot be treated similarly regarding the effects of age and education. In brief, neither youth nor higher education can protect the individual from the impairing effects of brain damage, and the severity, type, and location of brain damage become the principal determinants of the neuropsychological test results. It is obvious that if a neuropsychological test is principally determined by the condition of the brain, a major part of the variance cannot be determined by attribute variables.

These considerations suggest a gradient effect among psychological tests - the more sensitive a test is to brain impairment, the less it will be affected by attribute variables, and the less sensitive a test is to brain damage, the greater the likelihood that other variables will determine the results.

Similar studies were done using data from the Wechsler Adult Intelligence Scale (Reitan & Wolfson, 1996b, 1996d) and the results tended to show that a subtest's sensitivity to brain damage was a significant factor in determining the extent to which it was susceptible to age and education influences. Among brain-damaged groups, subtests that were sensitive to brain damage (such as Digit Symbol and Block Design) were relatively uninfluenced by age and education, whereas subtests that were relatively unaffected by brain damage (such as Information and Vocabulary) showed a significant relationship to age and, especially, to education.

The same research approach was used to investigate the effects of age and education in children. Reitan and Wolfson (1995b) found that both age and education had a strong effect on Neuropsychological Deficit Scale scores of non-brain-damaged control children. Brain-damaged children, however, not only showed no significant correlations between NDS scores and age and education, but, in addition, showed no significant improvement in NDS scores when comparing groups aged 9 through 11 versus 12 through 14 years of age. (This latter finding is of special concern, since it suggests that educational training had a negligible effect on the development of brain-based abilities in brain-damaged children. While non-brain-damaged children showed significant improvement in NDS scores in the two age ranges that were compared, brain-damaged children showed no significant improvement. The three years of educational training apparently were ineffectively applied in consideration of the developmental needs of children with brain damage!)

These findings also raised a question about the mandatory (inbuilt) age adjustment required to score the Wechsler Intelligence Scale for Children-Revised (as well as later versions). If brain-damaged children do not show significant improvement as they grow older, at least over a three-year period, what effect does this have on computation of Intelligence Quotients, since the norms are based on a presumption of improvement, as shown in normal children? The effect would be to produce lower IQ's for brain-damaged children as they grow older, resulting from the mere passage of time.

Considerable research, beginning mainly with Rourke (1975, 1985, 1989), has implicated impairment of brain functions among children with learning disabilities. A logical step, therefore, was to investigate the effects of age and education on neuropsychological test scores of children in this category. The results (Reitan & Wolfson, 1996a, 1996c) were similar to those found among children with definite diagnoses of brain damage or disease. These studies showed no significant relationships between age and education and NDS scores for children with learning disabilities. In addition, mean NDS scores of children aged 9 through 11 years versus 12 through 14 years were not significantly different. As with brain-damaged children, the effects of additional education appeared to have had no beneficial effect on the brain-based abilities of learning-disabled children.

This finding, as with brain-damaged children, raises a question about the computation of IQ's of children with learning disabilities. If the abilities of these children do not progress at the rate of normal children, but IQ computation presumes an age-based rate of development, the effect would be to produce progressively lower IQ's as the child ages. With respect to the age-old formula of intelligence based on the relationship between mental age and chronological age, only chronological age would remain as a variable, and the older the child becomes, the lower his IQ would be until adulthood.

These findings provide a basis for seriously questioning the mandatory scoring procedure of the WISC-R (and probably the WISC-III, as well), because an apparently unjustified adjustment, based solely on the results of normal children, is applied to the raw scores of children in whom age-related increments in performance may be limited. Is it of value to know how well the individual has actually performed, or is it sufficient to know how well he has performed on an adjusted basis compared to a normative sample? How many other tests may have this same problem? These results suggest that many other tests that make arbitrary adjustments of raw scores based on attribute variables should also be investigated for the actual meaning of their scoring procedures and the effect of such adjustments.

Reitan and Wolfson (1997, 1999) investigated the influences of age and education in cases of mild head injury. Two groups with mild head injury were studied. One group was composed of persons who had been routinely recruited for study, regardless of any symptoms or complaints, and a second group of equivalently mild head-injured persons was composed of patients who had eventually developed significant symptoms or complaints which had caused them to return for further evaluation and/or treatment.

Correlations of age and/or education with GNDS scores were not statistically significant in either group. Thus, the findings for persons with a mild head injury were essentially similar to the findings for persons with a more severe head injury, indicating that the neuropsychological consequences of the injury apparently had overridden the influences of age and education seen among non-brain-injured subjects. This finding must be interpreted in light of the sensitivity and specificity of the GNDS to the effects of brain damage and, obviously, may not apply to tests that are not sensitive to and specific for brain damage.

The consistency of the above findings tends to document their validity. In every study the effects of age and education on neuropsychological measurements were significant for non-brain-damaged persons, but were not significant for groups with brain impairment. This finding occurred in comparisons of controls versus brain-damaged adults, controls versus brain-damaged children, controls versus children with learning disabilities, and controls versus a group with mild head injuries. This finding appeared more prominently in tests that had known sensitivity to brain damage than in tests that tend to be resistive to the effects of brain damage.

The results of these studies are more completely reviewed in a paper by Reitan and Wolfson (1999). We should point out, however, that limited sample size and restrictions of the age and education ranges in these studies suggest that further research is needed. A very large sample might produce a statistically significant correlation, even though the magnitude of the coefficient is not increased nor the extent to which variance is explained. (The reader should be aware of this effect. For example, a correlation coefficient of .42 is required for significance at the .05 level with an N of 20, but a coefficient of only 0.19 is required if the N is 100.)

One may wonder how many other conditions there are, beyond impairment of brain functions, that might also diminish the effects of age and education on psychological and neuropsychological test scores, but this question has not yet been studied. It would seem clear, however, that adjustments based on age and education do not appear to be justified among persons or groups with brain impairment. Such adjustments, presently made on a routine and widespread basis, must be questioned, and additional research is needed to identify various categories of patients in whom the adjustments may or may not be justified.

Many studies over the years have shown definite relationships in adults, both in research and clinical contexts, between Wechsler scores and various aspects of brain damage (Reitan &

Wolfson, 1993), but the data used in these studies were based on scaled scores that were not adjusted for age or education. Studies of children, using scoring procedures that required age adjustment, have not shown similar findings. For example, there is not a single study in children that shows a significant difference between Verbal IQ and Performance IQ values, in cases with structural lateralization of cerebral damage (although many such studies have been published for adult groups).

Many variables and influences may be of significance in determining this circumstance, but transformation of scores for brain-damaged children on the basis of relationships that may not exist for them, but do exist for normal children, may have been a factor. Considering the "aging penalty" in scoring IQ's for brain-damaged and learning-disabled children (the older they become, the lower their IQ's become), as well as other brain-related variables such as lateralization effects, essentially the entire literature on intelligence and brain damage in children (which is mainly derived from Wechsler scores) may have been tainted.

If we postulate that brain damage has an adverse effect on neuropsychological test results in both adults and children, and that in this respect the categories represent different populations, should we not also entertain the hypothesis that attribute variables, such as age and education, might also differ in their effects on neuropsychological variables? Obviously, the universal and uncritical application of age and education levels, based on normal subjects, to brain-damaged persons and persons in all other diagnostic categories, has not considered this possibility. A helpful approach in dealing with this problem would be for normative data and scoring procedures to separate (rather than require) scores based on actual performances from scores that have been adjusted for age.

The Nature and Urgency of the Problem in Litigation

The following case illustrates one of many currently routine examples of the effects of transforming raw neuropsychological test scores to age- and education-adjusted scores.

The case involved a 51-year-old man with a 10th-grade education who had sustained a head injury four months prior to neuropsychological examination. He was suing for recovery of damages. Scores earned on the Halstead-Reitan tests that were administered are shown in Table 1.

According to the Heaton, Grant, and Matthews (HGM) T-scores, individual test scores fell in the normal range on six of the seven measures. Even though the Impairment Index fell in the impaired range, it was argued that it did so only because it was based on measures that were not adjusted for age and education, and thus was an inaccurate measure.

The NDS scores developed by Reitan and Wolfson (1988, 1993) were based on a review of many published studies and years of clinical application. Rather than depending on only one large sample, we were interested in taking advantage of consistency implicit in the many studies that have used the HRB. NDS scores for individual tests are represented on a four-point scale: 0 = a perfectly normal to superior performance; 1 = a score that is not exceptional, but still in the normal range; 2 = mild to moderate impairment; and 3 = severe impairment. In this case, the NDS scores showed an exactly opposite result from the HGM T-scores. Using NDS scores, six of the seven tests fell in the impaired range, and three of these scores represented severe impairment.

Since this case was involved in litigation, the plaintiff's neuropsychologist argued that the client was significantly impaired according to standards developed both clinically and in an extensive number of peer-reviewed publications; the neuropsychologist for the defense argued that outdated methods of evaluation were used by the plaintiff's neuropsychologist and that application of modern methods, taking age and education in consideration, showed that the client essentially fell in the normal range and had no significant impairment.

Obviously, striking disparities can result when conclusions are based upon conventional interpretation of the test data compared with the interpretation of the same data transformed according to the Heaton, Grant, and Matthews tables. Resolution of this disparity is clearly needed in clinical as well as forensic settings.

Review of Literature Concerned with Age and Education Transformations of HRB Scores

A number of critiques and rejoinders, as well as some empirical data, have been published about applying age and education corrections using the procedure proposed by Heaton et al. (1991), and the considerations raised in this paper cannot be evaluated fully without a review of these publications.

In 1991 Heaton, Grant, and Matthews published their book containing the procedures for transformation of raw scores for the HRB and other tests to T-scores adjusted for age and education. A review of this book by Fasteneau and Adams (1996) had a number of criticisms of the procedures used by Heaton, Grant, and Matthews.

Heaton et al. (1991) had created normative adjustments based upon ten age categories ranging from 20 years to 80 years and six education categories ranging from 6 to 18 years and more. Thus, adjustments were available for a total of 60 age versus education cells. Fasteneau and Adams (1996) noted that "a minimum of 15 subjects per cell is a rule of thumb for theoretical applications; 30 is a good minimum for norms in clinical decisions" (p. 446). They noted that the Heaton, Grant, and Matthews sample, based on "simple arithmetic yields a maximum of four people per subgroup, with as few as one person or even none at all representing at least some cells for some measures."

Heaton, Grant, and Matthews indicated that 244 men and 134 women were used to develop the normative samples. Thus, the total number of subjects used by Heaton, Grant, and Matthews appears to be insufficient. It should be noted, however, that Heaton, Grant, and Matthews pointed out that decisions of either a theoretical or clinical nature were not based upon the number of persons in each cell, but rather upon the overall trend represented for age and education.

Fasteneau and Adams (1996) also contended that distributions of age, education, and gender had not been demonstrated to be of significance for all of the tests used by Heaton, Grant, and Matthews. They said that these three variables were influential in only 14 of 54 measures, and considering the number of cells used for both men and women, the variables of age, education, and gender were not appropriate for the majority of the measures used. They also indicated that at the time of their review no one had presented a single substantive review or critique of the normative transformations proposed by Heaton, Grant, and Matthews.

Finally, Fasteneau and Adams (1996) questioned, "Can data generated on non-patient volunteers provide a useful benchmark to make statements about brain damage or dysfunction?" This comment relates to the fact that the normative data were based only on subjects without evidence of brain damage, whereas the interest of neuropsychologists is generally oriented toward identification or diagnosis of brain damage and its behavioral correlates.

Fasteneau and Adams (1996) offered pointed comments about the limited number of subjects in each cell, against which any individual might be compared on the basis of age and education. They cautioned that ". . . consumers should recognize the very significant limitations of the current volume and, if they use these norms, they should qualify their interpretations with professionally responsible statements such as those described in this review" (p. 448). This statement obviously was intended to emphasize the limited number of subjects in each cell that corresponded with the age and education of the person being evaluated.

Heaton, Matthews, Grant, and Avitable (1996) responded to the Fasteneau and Adams (1996) critique. They referred to the 486 normals, which included the group on which they based their normative data plus the group on which they based their validational data. They reported correlations of age and education with the Average Impairment Rating (AIR) for both the group of normal subjects (N = 486) and a group of brain-damaged subjects (N = 463). The correlation of the Average Impairment Rating with age was .71 for normals and .50 for brain-damaged subjects. The correlation of the AIR with education was .60 for the normals and .29 for the brain-damaged subjects.

A coefficient of .29 would account for only about 9% of the variance in the AIR distribution on the basis of age. It is apparent from these values that age and education, individually, had a significant effect on the AIR among normals, but this effect was substantially reduced in the group with brain damage. In fact, the authors state that age and education combined accounted for 64% of the AIR variance among normals, but for only 33.6% among brain-damaged subjects. The authors state that the diminished variance in test scores accounted for by age and education in the brain-damaged group is "to be expected due to their neurologic disorders, the severity of which may be unrelated to age and education" (p. 453).

However, the HGM procedure of adjusting HRB test results for age and education requires transformation to T-scores using tables based upon age and education. Heaton, Matthews, Grant, and Avitable (1996) transformed raw scores to age- and education-adjusted T-scores and found that differences between normal and brain-damaged groups were statistically significant, even though the T-score conversions had essentially eliminated the effects of age and education. The accuracy of classification of control subjects (specificity) and brain-damaged subjects (sensitivity) was determined for the AIR (which did not use age adjustments) and for the T-scores (which did use age and education adjustments). The T-score values were associated with a greater incidence of correct classifications of both controls and brain-damaged subjects: the controls rose from 82.5% to 85.8%, and the brain-damaged subjects rose from 66.1% to 75.8%.

These results might be considered quite impressive, except that analyses based upon General Neuropsychological Deficit Scale (GNDS) scores (without any corrections for age or education) achieved a 90% level of correct classifications for controls and a 92% level for correct classifications for brain-damaged subjects (Reitan & Wolfson, 1993). We would presume that the improved level of correct classifications using the GNDS resulted from the fact that the GNDS was designed to include all four methods of inference for detecting the presence of brain damage.

We should mention explicitly that Heaton, Matthews, Grant and Avitable (1996) argue that the number of subjects included in each cell, among the 60 cells used for age and education transformations, is not the pertinent point, because the T-scores were based on regression analyses using more than 350 subjects. The specific tables for each age and education category representing the 60 cells were designed for the convenience of transformations to T-scores in the individual case.

Heaton, Matthews, et al. (1996) felt that their age and education "corrections" achieved the following: (1) eliminated age and education effects from the test results, (2) showed similar error rates across age and education distributions, (3) showed similar results in clinical and normative groups, and (4) produced improved clinical accuracy rates compared with the AIR in correct classification of both controls (specificity) and brain-damaged persons (sensitivity). However, they also pointed out that "further research is needed to examine the performance of the T-scores with demographically diverse samples of subjects who have documented brain disorders" (p. 457). They felt that their age and education normative transformations, as compared with alternative technologies for interpreting performances on neuropsychological tests, "indicate that the norms do have significant advantages for neuropsychological clinical work and research" (p. 457).

Heaton, Ryan, Grant, and Matthews (1996) published a summarical position paper about the validity and significance of adjustments of neuropsychological test scores on the basis of age and education. In this paper they provided a detailed statement of the importance of attribute variables such as age, education, and gender as influences on neuropsychological test results in support of their investigations of these variables, and concluded that "the importance of such research cannot be overemphasized . . . and are critical to neuropsychology" (p. 143). As a final statement, these authors concluded that "without these [thorough and comprehensive norms], the interpretation of test scores will continue to rely more on educated guesswork than on science" (p. 160).

Heaton, Ryan et al. (1996) reached these conclusions without consideration of or reference to the many studies which have reported excellent validity findings *without the use of any age, education, or gender "corrections."* In an initial study using 50 controls matched in pairs with 50 brain-damaged subjects, not a single brain-damaged person earned a better Impairment Index than his matched control (Reitan, 1955). Over and beyond this finding, each of the seven tests on which the Impairment Index is currently based individually showed extreme probability levels in comparison of the two groups.

Wheeler, Burke, and Reitan (1963) analyzed HRB test results for four groups: Controls, Diffuse Cerebral Damage, Left Cerebral Damage, and Right Cerebral Damage using discriminant function analyses of neuropsychological test scores to determine classification of subjects to their correct groups. The results were as follows: Controls versus all categories of brain damage, 90%; Controls versus Left Cerebral Damage, 93.0%; Controls versus Right Cerebral Damage, 92.4%; Controls versus Diffuse Cerebral Damage, 98.8%; and Right Cerebral Damage versus Left Cerebral Damage, 92.9%. These results, based on HRB test scores alone, were the product of statistical analyses of the test scores, and *were not* adjusted for age or education. It would hardly seem that age and education adjustments were necessary to rescue the test data from a fate of "educated guesswork (rather) than science."

Wheeler et al. (1963) also determined the percentages of correct predictions of brain-damaged versus control subjects for each of 36 tests. In predicting membership in the control group, the 24-variable discriminant function, on which the above percentages were determined, was the most accurate, followed by the Impairment Index. Two measures based on critical flicker frequency were the least sensitive measures, ranking 35 and 36 among the 36 variables, and these test scores have been dropped from the HRB. Age and education were the next least accurate, ranking 33 and 34 among the 36 measures. This result provides further evidence of the limited significance of age and education as neuropsychological measures in comparison with the tests developed specifically to assess the integrity of brain functions.

In an earlier publication, Heaton, Grant, Anthony, and Lehman (1981) evaluated the accuracy of clinical conclusions about various aspects of brain damage as compared with the Average Impairment Rating (AIR). Ratings by two clinicians based on HRB protocols, yielded accuracy rates of 86.7% and 90.7% in differentiating between the presence and absence of brain damage, whereas only 75.3% were correctly predicted by the AIR. The clinicians were also more accurate in identifying lateralization of brain damage (left cerebral damage: clinicians, 71.1% and 80.6% versus 54.1% for the AIR; right cerebral damage: clinicians, 85.7% and 80.0% versus 72.3% for the AIR). Although age and education were included with the test scores in the evaluation process, the HRB scores were not adjusted or "corrected" on the basis of these variables. These results were not much different than the accuracy rates (noted above) that were achieved with T-scores adjusted for age and education, and they support the accuracy of clinical conclusions, as well as inter-judge reliability, in drawing inferences about brain damage (sensitivity) versus no brain damage (specificity) on the basis of raw HRB scores alone, without the use of age and education adjustments.

An noted previously, the General Neuropsychological Deficit Scale (GNDS) was developed to convert the HRB test data into a summary score that could be used directly in clinical evaluation. The GNDS is a direct reflection of raw test scores without any weighting for differential sensitivity among the 42 variables on which it is based or corrections for age and education. Published results (Reitan & Wolfson, 1988, 1993) were based on 169 persons with a variety of brain diseases and injuries and 41 controls who had been carefully examined medically for possible brain impairment or damage and for whom no such evidence was found. Many of the controls, however, were hospitalized and experiencing the stress that accompanies potentially serious illness.

A cutting point for the GNDS distributions was determined for differentiation of the two groups which classified 90% of the controls correctly (showing excellent specificity) and 92% of the brain-damaged persons correctly (showing excellent sensitivity). These values can be compared with the report of Heaton, Ryan, et al. (1996) based on T-scores on which age and education corrections had been made. As noted previously, these authors reported accuracy rates of 85% for controls and 75.8% for brain-damaged subjects. These results, especially for the brain-damaged group, were clearly below the rate achieved with the GNDS. From a clinical viewpoint, it is obviously of greater interest to recognize the effects of brain damage and to classify brain-damaged subjects correctly than to recognize normal scores and classify control subjects correctly.

Golden and van den Brock (1998) performed separate statistical analyses of Heaton, Grant, and Matthews T-scores and raw scores from the Halstead-Reitan Battery obtained in examination of 64 persons with focal cerebral lesions. They concluded that the "results certainly suggest that the patterns of scores in focal injuries generated by these [scoring] systems are indeed different," and that "one approach may not be appropriate for the other" (p. 163). These investigators found that raw scores showed evidence of more impairment than did scores based on the same data adjusted for age and education using the Heaton, Grant, and Matthews procedure.

Statistical analyses of data are entirely objective and impartial and are subject to inferential conclusions about chance probability. However, in practice, they can only serve as guidelines. The ultimate decisions (diagnoses) in the individual case are based on clinical judgments. Such clinical judgments may be viewed as less "scientific" than statistical analyses, but in actuality are ultimately required when evaluating the test results of individual clients. A valid question therefore concerns the validity of clinical conclusions about individuals. The ability to draw correct conclusions about individual persons is, in fact, "critical to neuropsychology" at a level well beyond corrections for age, education, and gender. The field was well aware of the critical importance of being able to reach correct clinical judgments long before issues of age and education were raised. In 1964 Reitan tested the accuracy of the Halstead-Reitan Battery, depending only on evaluation and interpretations of the test results without the use of age, education, or gender corrections, and the resulting findings would suggest that such adjustments are hardly "critical to neuropsychology" as claimed.

Based on neurological, neurosurgical, and neuropathological sources of information, we identified four individuals, each with either an intrinsic tumor, an extrinsic tumor, a focal traumatic lesion, or a focal vascular lesion, within each of the four lesion categories, which consisted of left anterior, left posterior, right anterior, and right posterior cerebral locations. We therefore had a total of 64 subjects with variously located cerebral lesions. Three additional groups were also included in the study: 16 subjects with diffuse cerebral damage or dysfunction due to cerebral vascular disease, 16 subjects with closed head injuries, and 16 subjects with multiple sclerosis. The inclusion of these groups provided a total of 112 patients. The results of the Halstead-Reitan Neuropsychological Battery were then reviewed to make "blind judgments" about the type and location (focal or diffuse) for each subject. The purpose, of course, was to determine the accuracy of inferences about brain lesions when based only on neuropsychological test data.

Using the neurological ratings as the criterion, there were 64 subjects with focal cerebral lesions. Of these subjects, 57 were classified correctly on the basis of psychological testing, and 7 were judged to have diffuse damage. In the group of 48 subjects with diffuse cerebral damage, 46 were classified correctly and 2 were judged to have focal lesions. The number of correct classifications on the basis of psychological test results for each of the lesion locations was as follows: left anterior, 9; left posterior, 11; right anterior, 7; and right posterior, 15. Thus, 42 of 64 patients were placed in their correct groups. Adding to this the correct classification of 46 of the 48 subjects with diffuse cerebral involvement, 88 of the 112 subjects were correctly classified.

For type of lesion, the 112 subjects were classified as follows on the basis of complete medical and surgical findings: intrinsic tumor, 16; extrinsic tumor, 16; cerebral vascular disease, 32 (16 focal, 16 diffuse); head injury, 32 (16 focal, 16 diffuse); and multiple sclerosis, 16. The number of correct classifications on the basis of psychological inferences was as follows: intrinsic tumor, 13; extrinsic tumor, 8; cerebral vascular disease, 28; head injury, 30; and multiple sclerosis, 15. Thus, 94 of the 112 patients were correctly classified according to type of lesion. Additionally, 13 of the 16 subjects with focal cerebral vascular disease were classified correctly, and 12 of these 13 were judged to have focal lesions. Of the 15 subjects with diffuse cerebral vascular disease who were correctly placed in this category, 14 were judged to have diffuse cerebral vascular disease. Thirty of the 32 head injury subjects had been identified correctly on the basis of their psychological test results, and 27 of these 30 had been classified correctly according to whether the lesion was focal or diffuse.

This degree of concurrence between neurological and neuropsychological ratings could scarcely have happened by chance. While adequate specificity statistics had been produced in earlier studies for brain-damaged and control groups, these results produced impressive specificity findings for both localization and type of brain damage based on HRB test scores. The results also confirmed that neuropsychological test results are differentially influenced by (1) focal and diffuse lesions, (2) which cerebral hemisphere is damaged, (3) frontal and nonfrontal lesions within the hemisphere involved, (4) intrinsic tumors, extrinsic tumors, cerebral vascular lesions, head injuries, and multiple sclerosis, (5) focal occlusion as compared with generalized insufficiency in cerebral vascular disease, and (6) focal as compared with diffuse damage from head injuries.

These findings, as noted, were based on test results that had not been adjusted for age or education, although information about these variables was known for each person at the time the ratings were made. Evidence regarding the extent to which neuropsychological findings relate to the differential aspects of brain pathology is particularly important in the forensic area inasmuch as critical questions so often concern the questions about possible brain damage.

The Use of Age and Education in Clinical Evaluations

The reader may be interested in the manner in which age and education data are used in clinical evaluation of the individual person, since such data frequently do serve a useful role. Many examples have been provided in addition to a review of published studies (Reitan, 1979). It should be recognized that we have routinely found these variables to be related to neuropsychological test scores among non-brain-damaged persons, and have never proposed that brain-damaged subjects are entirely immune to these effects.

Our position in this regard, as stated previously, hinges on the sensitivity and specificity of tests to brain damage. Among persons with brain damage, the more sensitive and specific the test is to brain damage, the less the available variance that may be determined by other variables such as age and education. After all, sensitivity and specificity are critical requirements of neuropsychological tests. If neuropsychological tests were not sensitive to brain damage, but were influenced principally by other factors, they could hardly serve as indicators of brain status.

Thus, we propose that an inverse relationship exists between the sensitivity and specificity of a test to brain damage and the extent to which age and education may determine the results of that test among persons with brain damage. This proposal carries with it many complexities, considering the different ways in which subcategories of brain damage (left versus right cerebral lesions, for example) may affect test scores differentially. In general, however, neuropsychological interpretation of the HRB in the individual case provides relevant conclusions that clarify considerations of this type.

Basically, the procedure in interpretation of individual protocols is first to determine whether evidence of brain damage is present. (Unfortunately, many "flexible" batteries cannot provide an answer to this question.) If brain damage is present, one can expect the tests most sensitive to brain damage to be impaired, regardless of the age or education of the client, and one would proceed to consider the full range of additional dimensions included under the general rubric of brain damage.

The effects of brain damage are obviously quite variable, depending upon location, type, status, etc. of damage caused by the injury or disease, but these effects occur essentially regardless of age or education. It should be noted that tests which are not specifically sensitive to brain damage may be relatively spared, and may still largely reflect an individual's age and educational status. Other tests may occupy an intermediate position, still retaining age and/or education effects even though they have been influenced to a degree by brain damage. Clearly, as with all clinical evaluations, the knowledge and competence of the clinician are major factors in understanding the complexities implicit in the test data. Unless the clinician is informed about the ways in which age and education variables affect the test results, and which tests are more and less sensitive to brain damage, the above recommendations can hardly be implemented.

It is beyond the scope of this paper to attempt a full exposition of the processes by which valid clinical conclusions are reached in evaluation of neuropsychological test data (see Reitan & Wolfson, 1986b, 1988, 1993). With respect to age, education, and brain damage, the general rule is to look for definitive indications of brain damage which, if found, would stand as valid indications regardless of the subject's age or education. If no indications of brain damage are found, the usual effects of age and education should be considered as appropriate.

Our usual procedure is to review the raw-score test data initially and to record written conclusions at least in brief form, and then turn to any available history and/or neurological diagnostic information for correlation with the test results. The sequence with which these categories of information are reviewed varies among individual neuropsychologists. If knowledge of the presence of brain damage is known before the neuropsychological test records are reviewed, such information may provide a basis for expecting the effects of the damage to have a distinct influence on any effects of education or age.

Neuropsychologists were aware of the importance of understanding the differential and interactional effects of age and education variables long before the current issues about corrections or adjustments arose, and many studies were performed to evaluate the effects of age and education with relation to psychological and neuropsychological test results. Some of this research has been summarized by Reitan (1970a) and Reitan and Wolfson (1986a). Implementation of the plan for clinical interpretation, briefly described above, is greatly facilitated by knowledge of which neuropsychological tests are most sensitive to brain damage, which tests are most affected among normals by aging, and which tests are relatively unaffected by brain damage but are influenced by age and education. Clinical experience involving many brain-damaged and non-brain-damaged persons is essential in learning to understand the interactions of test results, and this type of knowledge can only be gained in any detail by using a constant basic test battery.

There is nothing at this point to restrict comparison of the individual's test results with those of persons either with or without brain damage. In fact, such comparisons represent the fundamental basis for conclusions about the presence or absence of brain pathology. The point of this paper is not to limit or restrict such comparisons; our concern is to make such comparisons with a full and accurate knowledge of the groundwork and procedures on which the comparisons are based. If, for example, a subject's occupation was known to have a specific and relatively invariant influence on the score obtained on a particular test, the neuropsychologist would want to consider occupation as a variable and the possible differential significance of the particular occupation among persons with normal as opposed to damaged brains. Obviously, consideration of the subject's occupation would lose interest if research results had shown that the particular occupation lost its significance regarding the test score if brain damage was present.

Conceptual Considerations Concerning Age and Education Transformations

Table 2 of Heaton, Matthews, et al. (1996) makes it quite clear that the raw Average Impairment Rating (AIR) scores of the least educated and older subgroups were lower than for any other subgroup, and the presumption was that this was due to limited education and to older age. Thus, their rationale would be that raw scores in these subgroups were in need of "correction." However, there are many reasons for delving a little deeper into this explanation.

Corrections for age and education, even among supposedly normal persons, may represent a significant conceptual error, at least among older and less educated persons. Declining performances with advancing age among normals may, in fact, represent a correction for brain impairment. Neurological studies, completely separate from neuropsychological testing, have shown unequivocal findings of structural, physiological, and neurobehavioral changes that occur in the aging process. Many of these studies have been reviewed by Reitan and Wolfson (1992).

For example, findings on autopsy of presumably normal persons show many changes of the type seen in Alzheimer's disease, not only among subjects in North America, Great Britain, and European countries, but also in Japan. Matsuyama and Nakamura (1977) reported on their examination of 617 brains in Tokyo. These investigators deliberately excluded brains of patients who had evidence of psychosis or any conditions known to be predisposed to the types of pathological changes seen in Alzheimer's disease. Thus, these patients presumably represented a sample of subjects who might manifest the effects of normal aging but not brain disease.

The results of this study indicated a strong association between advancing age and the occurrence of neuropathological changes, including neurofibrillary tangles and senile plaques, which are the major pathological hallmarks of Alzheimer's disease. Such changes were present in approximately 50% of cases aged from 50-59, in more than 80% of cases aged 60-69, and in almost every patient over age 70. Changes of this kind represent a difference in *degree* of pathology in aging and Alzheimer's disease, but no apparent difference in *kind*.

Obviously, it is difficult to relate autopsy findings to the performance capabilities of living persons. However, it has been possible to compare the type and degree of neuropsychological deficits found with younger brain-damaged persons and the neuropsychological test results among older neurologically-normal persons. Reed and Reitan (1963) found statistically significant evidence that the neuropsychological tests that were most sensitive to brain damage in younger persons were also the tests that most strikingly reflected intellectual and cognitive impairment among older persons, even though this latter group was composed of persons who appeared, on clinical neurological examination and by history, to be aging normally. Thus, both the pathological changes of the brain and the neuropsychological findings among supposedly normally aging people are similar to findings of groups with definite brain damage.

The adjustments made among the elderly by the Heaton, Grant, and Matthews procedures are aimed at eliminating the effects of age and education differences. In effect, the "corrections" eliminate the neuropsychological impairment seen among older persons because this impairment is presumably "normal," even though independent evidence indicates that it is associated with pathological brain changes. It hardly seems surprising, then, that very poor scores are classified as normal when the HGM "corrections" are made. This effect is illustrated by the fact that a 55-year-old man with 17 years of education can still make 80 errors on the Category Test and be classified by the HGM transformation as having performed normally. The validity of a result of this kind is difficult to understand, considering that Reitan (1955) found that a control group made an average of 32 errors and a brain-damaged group made an average of 62 errors!

Thus, while HGM corrections are intended to eliminate the effect of age and education differences, thus providing for comparable T-scores across each distribution, it would seem that they are, in effect, at least partially eliminating the effects of brain pathology on neuropsychological test scores. This procedure would have its major effect in producing normal scores among brain-damaged persons, since the normative data on which the adjustments were based were derived, in part, from older persons and less-well educated adults whose poorer raw scores were very probably determined, in part, by brain impairment.

In summary, an abundance of evidence has linked older age to brain pathology, even among older persons without a clinical diagnosis such as Alzheimer's disease. There can hardly be any disagreement that these neuropathological changes are the bases for neuropsychological impairment. In fact, the authors of the special article on neuropsychological assessment (Neurology, 1996) state that aging negatively affects intellectual abilities, complex attentional processes, some aspects of memory, psychomotor speed, assessing word knowledge, visuospatial skills, some forms of reasoning, and complex motor problem-solving. Presumably, some factors related to the impairment may derive from limited sensory or motor functions, but a major component of the cognitive decline is due to gradually progressive brain impairment.

The fact that a substantial part of the need for corrections for education came from persons with lesser education calls for a similar explanation. Obviously, there are reasons that these people achieved a lesser education, and some of the reasons have to do with circumstances that would have no connection with the adequacy of brain functions. It is likely, however, that the lower educational level of these adults was due, in part, to limitations of brain functions, which in turn were expressed as lowered intelligence, lessened academic aptitude, and poorer test scores.

Heaton et al. (1991) do not seem to have appreciated the relationship of advanced age and limited education to impairment of brain functions, even though their sample had been determined to be neurologically normal. In the first paragraph of their book they cite the usefulness of neuropsychological tests in evaluation of brain disorders, but then state, "However, neuropsychological tests are not exclusively sensitive to brain pathology. It has long been recognized that performances on most of these instruments are strongly related to age and education . . ." (p. 3). This statement appears to deny the distinct likelihood that age and education, particularly at the extremes of the two distributions (older age and lesser education), almost certainly, in many cases, have a definite relationship to the biological status and adequacy of brain functions.

What, then, is the result of developing raw score/T-score transformations based on age and education? The result is to produce normal T-scores particularly for persons with older age and lesser education. In effect, the result is to produce normal scores for the parts of the distribution in which raw scores have been poorest, and poorest, at least in part, because of brain impairment. What, then, would be likely to happen if such adjustments were made in a group of persons who all have brain impairment? The expected result would be that they would score better when the adjustments were made (T-scores) than when no such adjustments had been made (raw scores. The above rationale is consistent with the findings reported by Reitan and Wolfson (1995a) in their study of relationships between age and education and GNDS scores. Among their control (non-brain-damaged) subjects, the correlation between age and education differed sharply for the older and younger halves of the distribution. The older half had a coefficient of 0.68 (showing that GNDS scores were poorer as age increased), whereas the younger half had a coefficient of only 0.10.

These adverse effects of aging, as shown on the GNDS, were entirely consistent with the expected neuropathological changes that develop with advancing age, even among persons classified as controls. Age adjustments in this group, done to eliminate age effects, are in large part equivalent to eliminating the effects of brain impairment. The very reason that neuropsychological tests are used is to identify (not eliminate) the effects of brain impairment. Thus, for the older segment of the so-called "normal" population, the transformations proposed by Heaton et al. (1991) appear to be conceptually flawed, inasmuch as aging is a variable that causes subtle brain damage in many so-called "normal" people.

The same line of reasoning applies to the segment of the "normal" population with limited education, even though neuropathological studies have not focused on this group as they have for the aged. As noted previously, multiple factors may be responsible for limited educational achievement, but adequacy of brain functions, resulting in limitations of intelligence and academic aptitude, is surely among them. For example, on brain-sensitive tests, would one normally expect persons with a low educational level to perform as well as persons with more education? Reitan and Wolfson (1995a), when dividing their control group at the 50% percentile for educational level, found that the lower 50% had a correlation of -0.56 ($p<.01$) between education and GNDS (the lower the education, the poorer the GNDS score), whereas the upper 50% had a correlation of 0.27, a value that did not reach significance considering the size of the group.

Undoubtedly, the overall role of age and education among neurologically-normal groups is hardly to be explained by postulating that subtle impairment of brain functions is the only (or perhaps, even major) factor for elderly persons or for the lesser educated. It would appear very likely, however, that brain impairment is one of the relevant variables, and the above considerations were presented in order to recognize the complexity of the entire situation. Can comparable procedures for effecting age and education "corrections" be used across the entire span of the age and education distributions, even among "normals?" Can age and education adjustments, aimed at removing the effects of these variables, be acceptable when they actually remove, at least in part, the effects of brain impairment - the very condition that the tests were designed to detect?

In summary, when one adjusts neuropsychological test scores for age, one adjusts the test scores for all of the additional variables that are significantly associated with age. These additional factors will be related to neuropsychological test scores to a varying degree. Some variables, such as reduced physical strength, reduced sexual drive, greying of hair, etc. may have little or no relationship to test scores, and therefore the adjustments of test scores for age will have little if any consequence. Motor speed and dexterity, reaction time, reasoning and logical analysis, and other neuropsychological functions may be more adversely affected by age, and poor scores on tests that require these skills would be changed more strikingly by age adjustments in the direction of normality. Brain impairment, which is generally but variably more pronounced with aging, could be a major factor, and tests that are particularly sensitive to brain impairment could be done very poorly and still be considered normal after age adjustments had been made.

Since both older age and limited education appear to be associated with variables that produce poorer raw scores on neuropsychological tests (such as pathological brain changes), the extent to which these associated variables produce lower scores will determine the degree of adjustment necessary to bring the performances up to a mean T-score of 50. In other words, the more sensitive the test is to these associated variables, the poorer the raw score can be and still be classified as normal after the adjustment is made.

Results of a Current Study

The present investigation was concerned with a direct comparison of transformations (called "corrections" by Heaton et al. [1991]) of raw scores from the Halstead-Reitan Neuropsychological Test Battery for Adults to T-scores by using the HGM tables and NDS scores by using the table provided by Reitan and Wolfson (1993). To summarize the information given previously, the HGM procedure first transforms raw scores into scaled scores, and the scaled scores are used to obtain T-scores (with a mean of 50 and a standard deviation of 10) depending upon the gender, age, and education of the subject.

The Reitan and Wolfson procedure makes no adjustment for age or education, but merely uses the raw score to determine the corresponding NDS score. NDS scores range from 0 (perfectly normal) to 3 (severely impaired), and were devised originally to provide information about the clinical significance of raw scores on each of the 42 variables in the HRB and to place the score for each test on a standard scale so that the scores could be added and thus provide a total score for the entire battery (the General Neuropsychological Deficit Scale score).

Purpose. Although there are many ways in which comparisons of the two methods of transforming raw scores could be evaluated, it seemed of value initially to use a procedure that utilized the fact that each method identified a cutting-point for separation of "normal" and "impaired" scores (HGM method: T-scores of 39 and less; Reitan and Wolfson method: NDS scores of 2 and 3). The purpose of this study was to determine and compare the frequency with which "normal" and "impaired" scores occurred in a sample of non-brain-damaged controls and a sample of persons with diffuse cerebral vascular disease.

Description of Groups. In order to avoid any possibility that group composition might have been influenced by the purposes of this study, the groups compared were ones that had been used in a prior study done years before the question of normative transformations was even proposed (Reitan, 1970b). The control group was composed of 26 men who were functioning normally in their occupations and had no reports of illness or injuries they felt might have compromised their intellectual and cognitive competence. The group had a mean age of 54.00 years (SD, 4.22) and a mean education of 14.73 years (SD, 2.75).

The group with diffuse cerebral vascular disease consisted of 26 men who had a mean age of 53.54 years (SD, 6.28) and a mean education of 11.00 years (SD, 3.03). The difference in age distributions between the two groups was negligible (t ratio, 0.27; $p<.70$). Education differences, however, were significant (t ratio, 4.55; $p<.001$).

It must be pointed out, as was done in the original report, that a conclusion of diffuse, generalized, or multifocal cerebral vascular disease is relatively difficult to reach as a firm differential neurological diagnosis. For this reason, detailed results of the clinical neurological examination, specialized neurological diagnostic procedures, frequency of complaints of a neurological nature reported by the patients, and associated conditions or diagnoses were included in the description of the group (Reitan, 1970b, pp. 159-160). The diagnosis of diffuse cerebral vascular disease was not dependent on any single one of these sources of information, but instead represented the definitive conclusion of the examining neurologist.

Procedure. In accordance with the purpose of this study, the raw scores for nine variables were transformed to both HGM T-scores and to Reitan and Wolfson NDS scores for each group. The number of scores falling in the impaired and normal ranges was determined for both groups according to each method. Differences in the distributions using the two methods were tested for significance using a chi-square test based on a four-celled table. These results permitted an inference of whether the distribution of normal and impaired scores, produced by the two methods, differed beyond chance expectancy.

Results. Table 2 presents means and standard deviations for the nine tests used in this study, calculated from raw scores, from T-scores derived from the HGM transformations and from Reitan and Wolfson's NDS scores. We have not included statistical comparisons of the groups, since this information was included in the original report (Reitan, 1970b). However, the controls consistently performed better than the group with cerebral vascular disease at statistically significant levels.

Our present purpose was to determine the frequencies with which normal and impaired scores were obtained when using the HGM and RW methods, and this information is presented in Table 3.

The control group, which had no evidence of brain involvement, consistently showed a greater frequency of normal scores than impaired scores, regardless of the transformation method that was used. Overall, the HGM method classified 87% of these scores in the normal range (T-scores of 40 or higher) and the RW method classified 76% in the normal range (NDS scores of 0 or 1). Thus, more of the controls showed no impairment using the HGM method than the RW method. A chi-square test of these distributions yielded a value of 15.39 ($p<.001$).

Conversely, the RW method consistently classified persons with diffuse cerebral vascular disease as having impaired scores more frequently than did the HGM method. The RW method classified 82% of the scores in the CVD group as being impaired, whereas the HGM method classified 63% as being impaired. The chi-squared comparison of these differences yielded a value of 16.74 ($p<.001$).

It is apparent from these results that the two methods, using the same raw-score data, differ significantly in the frequency with which the transformations identify impaired and normal scores. The HGM method identifies more scores as normal, whereas the RW method identifies more scores as impaired. The consistency of this observation is indicated by the finding that in the control group the HGM method showed a greater percentage of subjects earning normal scores on 8 of the 9 variables (with one tie) than did the RW method; in the CVD group, all nine tests showed a greater frequency of impairment using the RW method than when using the HGM method. There would seem to be no doubt that the two methods produce results that differ in their clinical significance.

Discussion. Some controls will, of course, produce some scores in the impaired range, and some brain-damaged subjects will produce some scores in the normal range. However, there is relatively little available information to use as a guideline for expectation of the frequency of impaired scores for persons who qualify as normal controls versus expectation of the frequency of normal scores for persons with evidence of brain damage (and even less information for persons with diffuse cerebral vascular disease). It might be noted, however, that there is a long historical record in clinical interpretation of indicating that control subjects may have Impairment Indexes of up to 0.3 (30% of the tests in the impaired range), with an Impairment Index of 0.4 (40% of the tests in the impaired range representing the borderline range.

On this basis, it might seem that as few as 13% of the scores in the impaired range among control subjects, as found with the Heaton, Grant, and Matthews transformations, may represent an underestimate of the number of "impaired" scores, especially considering the age of the subjects in this study. The standard interpretation of the Impairment Index, however, is based on test results that were not adjusted for age and education, and the age and education corrections of the HGM method may be producing more accurate estimates of improved normality and functional capabilities. The incidence of 82% of the patients with cerebral vascular disease earning impaired scores when using the Reitan and Wolfson criteria hardly seems surprising, and the finding that only 63% of the scores for the CVD group fell in the impaired range when using the HGM method may well be an underestimate.

Certain implications of the present findings, based on the groups studied, are quite clear. Using the HGM method to transform HRB raw scores results in generation of more scores in the normal range for controls than did using the Reitan-Wolfson method. Conversely, the Reitan-Wolfson method produced more scores in the impaired range for a group with diffuse cerebral vascular disease.

If these results are confirmed and are determined to represent a general conclusion, the findings will have distinct and definite significance not only in clinical evaluations, but also for forensic interpretations. HGM transformations of data will tend to indicate that a client is unimpaired, whereas Reitan-Wolfson transformations of data will tend to emphasize impairment. If a plaintiff is seeking compensation for brain impairment, use of HGM transformations would be of benefit for the defense, inasmuch as indications of impairment would tend to be minimized. However, in the same case, the indications of impairment might be emphasized through use of the Reitan-Wolfson transformations because the client's age and education were not represented in the scores. Obviously, if one method of transformation tilts the interpretation either toward normality or toward impairment more significantly than another method of transformation, resolution of the conflict is needed.

It must be pointed out explicitly that the empirical results reported in the study represented a limited investigation of this complex issue. The basic purpose of this paper was to stimulate further investigation, and this should be quite possible, considering the wealth of data on the Halstead-Reitan Battery that is available in the field. Not only the effect on test scores (as studied in this paper), but the effect on classifications of groups of brain-damaged and control subjects, according to age, type of lesion, lateralization and location of lesion, etc. must also be investigated. The urgency of such investigations is implicit, inasmuch as clinical conclusions are routinely being drawn daily, using transformations of raw scores, based on the age and education of non-brain-damaged subjects, applied to brain-damaged subjects, with limited empirically-derived knowledge of the accuracy of these transformations.

The need for such additional investigations is evident both for clinical interpretation and for use of neuropsychological data in a forensic setting. Many neuropsychological tests, over and beyond those referred to in this paper, may also be influenced by the age and education of the particular person being evaluated, and attenuation of age and education effects may be differentially affected in accordance with the sensitivity of the test to brain damage generally and, more specifically, to variables such as diffuse versus focal damage, the particular location of damage, and the status of the brain disease or injury. In addition, the interaction of age and education with variables describing brain involvement may differ depending upon whether the individual is young or old, well educated or poorly educated. Real-life variables are complex and intertwined. While the overall aim of science is to gain valid generalizations, expert and valid conclusions as clinicians and forensic experts focus on the individual and we must continually be aware of the uniqueness of every individual person.

References

Assessment: Neuropsychological testing of adults. Considerations for neurologists. 1996. Report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. *Neurology, 47*, 592-599.

Fasteneau, P.S., & Adams, K.M. 1996. Heaton, Grant, and Matthew's Comprehensive Norms: An overzealous attempt. *Journal of Clinical and Experimental Neuropsychology, 18,* 444-448.

Golden, C.J., & van den Brock, A. 1998. Potential impact of age- and education-corrected scores with HRNB scores patterns in participants with focal brain injury. *Archives of Clinical Neuropsychology, 13,* 683-694.

Heaton, R.K., Grant, I., Anthony, W.Z., & Lehman, R.A. 1981. A comparison of clinical and automated interpretation of the Halstead-Reitan Battery. *Journal of Clinical Neuropsychology, 3,* 121-141.

Heaton, R.K., Grant, I., & Matthews, C.G. 1991. *Comprehensive norms for an expanded Halstead-Reitan Battery*. Odessa, FL: Psychological Assessment Resources.

Heaton, R.K., Matthews, C.G., Grant, I., & Avitable, N. 1996. Demographic corrections with comprehensive norms: An overzealous attempt or a good start? *Journal of Clinical and Experimental Neuropsychology, 18,* 449-458.

Heaton, R.K., Ryan, L., Grant, I., & Matthews, C.G. 1996. Demographic influences on neuropsychological test performance. In I. Grant & K. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders.* (2nd ed.) (pp. 141-163). New York: Oxford University Press.

Jarvis, P.E., & Barth, J.T. 1994. *The Halstead-Reitan Neuropsychological Battery: A Guide to Interpretation and Clinical Application.* Odessa, FL: Psychological Assessment Resources.

Matsuyama, H., & Nakamura, S. 1977. Senile changes in the brain in the Japanese: Incidence of Alzheimer neurofibrillary and senile plaques. In R. Katzman, R.D. Terry, & K.L. Bick (Eds.), *Alzheimer's disease: Senile dementia and related disorders. Aging, Vol. 7* (pp. 287-297). New York: Raven Press.

Oestreicher, J.M., & O'Donnell, J.P. 1995. Validation of the General Neuropsychological Deficit Scale with nondisabled, learning-disabled, and head-injured young adults. *Archives of Clinical Neuropsychology, 10,* 185-191.

Reed, H.B.C., & Reitan, R.M. 1963. A comparison of the effects of the normal aging process with the effects of organic brain damage on adaptive abilities. *Journal of Gerontology, 18*, 177-179.

Reitan, R.M. 1955. An investigation of the validity of Halstead's measures of biological intelligence. *Archives of Neurology and Psychiatry, 73*, 28-35.

Reitan, R.M. 1964. Psychological deficits resulting from cerebral lesions in man. In J.M. Warren & K.A. Akert (Eds.), *The frontal granular cortex and behavior*. New York: McGraw-Hill.

Reitan, R.M. 1970a. Measurement of psychological changes in aging. *Duke University Council on Aging and Human Development*, Proceedings of Seminars.

Reitan, R.M. 1970b. Objective behavioral assessment in diagnosis and prediction. In A.L. Benton (Ed.), *Behavioral change in cerebrovascular disease* (pp. 155-165). New York: Medical Department, Harper & Row.

Reitan, R.M. 1979. *Neuropsychology and aging*. Tucson, AZ: Neuropsychology Press.

Reitan, R.M., & Wolfson, D. 1986a. The Halstead-Reitan Neuropsychological Test Battery and aging. In T.L. Brink (Ed.), *Clinical gerontology: A guide to assessment and intervention* (pp. 39-62). New York: Haworth.

Reitan, R.M., & Wolfson. D. 1986b. *Traumatic brain injury. Vol. I. Pathophysiology and neuropsychological evaluation*. Tucson, AZ: Neuropsychology Press.

Reitan, R.M., & Wolfson, D. 1988. *Traumatic brain injury. Vol. II. Recovery and rehabilitation*. Tucson, AZ: Neuropsychology Press.

Reitan, R.M., & Wolfson, D. 1992. *Neuroanatomy and neuropathology: A clinical guide for neuropsychologists* (2nd ed.). Tucson, AZ: Neuropsychology Press.

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

Reitan, R.M., & Wolfson, D. 1995a. Influence of age and education on neuropsychological test performances of older children. *Child Neuropsychology, 1,* 165-169.

Reitan, R.M., & Wolfson, D. 1995b. Influence of age and education on neuropsychological test results. *The Clinical Neuropsychologist, 9,* 151-158.

Reitan, R.M., & Wolfson, D. 1996a. Can WISC-R IQ values be computed validly for learning-disabled children? *Applied Neuropsychology, 3*, 15-20.

Reitan, R.M., & Wolfson, D. 1996b. Differential relationships of age and education to WAIS subtest scores among brain-damaged and control groups. *Archives of Clinical Neuropsychology, 11*, 303-311.

Reitan, R.M., & Wolfson, D. 1996c. The diminished effect of age and education on neuropsychological performances of learning-disabled children. *Child Neuropsychology, 2,* 11-16.

Reitan, R.M., & Wolfson, D. 1996d. Relationships of age and education to Wechsler Adult Intelligence Scale IQ values in brain-damaged and non-brain-damaged groups. *The Clinical Neuropsychologist, 10,* 293-304.

Reitan, R.M., & Wolfson, D. 1997. The influence of age and education on neuropsychological performances of persons with mild head injuries. *Applied Neuropsychology, 4,* 16-33.

Reitan, R.M., & Wolfson, D. 1999. The influence of age and education on neuropsychological performances of persons with mild traumatic brain injuries. In M.J. Raymond, T.L. Bennett, L.C. Hartlage, & C.M. Cullum (Eds.), *Mild traumatic brain injury.* Austin: Pro-ed.

Rojas, D.C., & Bennett, T.L. 1995. Single versus composite score discriminative validity with the Halstead-Reitan Battery and the Stroop Test in mild head injury. *Archives of Clinical Neuropsychology, 2,* 101-110.

Rourke, B.P. 1975. Brain-behavior relationships in children with learning disabilities: A research program. *American Psychologist, 30,* 911-920.

Rourke, B.P. (Ed.). 1985. *Neuropsychology of learning disabilities: Essentials of subtype analysis.* New York: Guilford Press.

Rourke, B.P. 1989. *Nonverbal learning disabilities: The syndrome and the model.* New York: Guilford Press

Wheeler, L., Burke, C.J., & Reitan, R.M. 1963. An application of discriminant functions to the problem of predicting brain damage using behavioral variables. *Perceptual and Motor Skills, [Monograph supplement], 16*, 417-440.

Table 1. Test results for a 51-year-old man in which the Heaton, Grant, and Matthews T-scores and the Reitan and Wolfson NDS scores disagree.

Test	Raw Score	HGM T-Score	R & W NDS
Category	82	42	3
Tactual Performance Test			
Total Time	21.7	44	2
Memory	6	42	1
Localization	2	46	3
Speech-sounds Perception Test	20	33	3
Rhythm Test	23	43	2
Finger Tapping - Preferred Hand	49	46	2
Impairment Index	0.9	29	3

Table 2. Means and standard deviations for a control group and a group with diffuse cerebral vascular disease, representing raw scores, Heaton, Grant, and Matthews transformations, and Reitan and Wolfson transformations.

Group		Category	**TPT Time**	**TPT Mem**	**TPT Loc**	Rhythm	**Speech-Sounds**	Tapping	Trails A	Trails B
Raw Scores										
Controls	M	56.60	14.30	6.96	3.15	27.23	4.15	52.73	33.08	74.77
	SD	29.05	3.78	1.18	1.87	1.99	2.98	3.96	10.15	23.63
CVD	M	85.00	33.05	4.58	1.69	19.27	17.38	38.27	84.15	284.42
	SD	36.81	13.56	2.67	1.89	4.39	13.99	12.48	67.53	268.45
HGM T-scores										
Controls	M	44.35	52.50	45.00	48.58	55.08	53.35	52.38	48.00	51.31
	SD	11.48	7.52	9.38	7.35	9.99	8.42	6.73	9.20	7.67
CVD	M	37.34	35.92	35.42	42.50	39.00	37.04	36.46	34.54	34.65
	SD	11.06	13.79	14.66	9.25	14.00	7.03	11.67	10.00	14.76
RW NDS Scores										
Controls	M	1.73	1.31	0.92	1.27	0.69	0.15	1.00	1.00	1.04
	SD	1.04	0.79	0.84	0.92	0.76	0.46	0.69	0.80	0.96
CVD	M	2.54	2.50	1.88	2.65	2.12	2.00	2.35	2.38	2.50
	SD	0.71	0.71	1.14	0.75	1.14	0.94	0.94	0.80	0.99

Table 3. Percentage of scores for 26 persons which were classified as Normal or Impaired using the Heaton, Grant, and Matthews transformations and the Reitan and Wolfson transformations.

		Category %	**TPT Time %**	**TPT Mem %**	**TPT Loc %**	Rhythm %	**Speech-sounds %**	Tapping %	Trails A %	Trails B %	Total %
HGM Transformations											
Controls	Normal	69	96	69	92	92	100	96	77	92	87
	Impaired	31	4	31	8	8	0	4	23	8	13
RW Transformations											
Controls	Normal	65	65	69	85	85	96	77	77	65	76
	Impaired	35	35	31	15	15	4	23	23	35	24
HGM Transformations											
CVD	Normal	50	42	35	46	38	31	42	31	19	37
	Impaired	50	58	65	54	62	69	58	69	81	63
RW Transformations											
CVD	Normal	12	12	27	8	27	27	23	12	12	18
	Impaired	88	88	73	92	73	73	77	88	88	82