Applied Neuropsychology
2005, Vol. 12, No. 4, 181–189

The Effect of Age and Education Transformations on Neuropsychological Test Scores of Persons With Diffuse or Bilateral Brain Damage[1]

Ralph M. Reitan and Deborah Wolfson

Reitan Neuropsychology Laboratory, Tucson, Arizona, USA

Heaton, Grant, and Matthews (1991) published procedures for adjusting raw scores on various neuropsychological tests according to the individual's age and education. Despite rather widespread use of these score conversions in both clinical work and research publications, there have been very few investigations to evaluate the accuracy or limitations of these score transformations. This study was based on 52 persons with brain damage who had a range of verified brain injuries or disease. Raw scores were transformed according to the Heaton, Grant, and Matthews (HGM) procedure and also according to Reitan and Wolfson's Neuropsychological Deficit Scale (RW NDS) score procedure. The latter method is a straightforward transformation of raw scores and does not adjust for age or education. The number of normal (vs. impaired) scores was 1.74 times greater using the Heaton, Grant, and Matthews method than when using the Reitan and Wolfson method, yielding a χ^2 value of 32.66 ($p <$ .001). These results suggest that impairment in this brain-damaged group was identified less often when using the Heaton, Grant, and Matthews procedure than when using a method that represented the raw scores directly. The clinical and forensic implications of these findings are discussed. More research on this issue is needed.

Key words: age effects, education effects, score transformations, Heaton, Grant, and Matthews procedure, Reitan and Wolfson procedure, transformation validity, Halstead–Reitan Battery

In 1991 Heaton, Grant, and Matthews (HGM) published procedures for adjusting raw scores on various neuropsychological tests according to the individual's age and education. These variables were presumed to affect the individual's test scores, and the purpose of the adjustment was to reflect those effects. Although the procedure was developed using data from a group of neurologically normal persons, in practice the procedure seems to be applied mainly to persons with brain impairment. Adjusting test scores for age and education appears to be widespread in both clinical and forensic settings. For example, Jarvis and Barth (1993) presented HGM converted scores for each illustrative case in their book on interpretation of results from the Halstead–Reitan Battery.

Although there is a rather general adoption of the HGM score-transformation procedure, very little basic investigative work has been done to evaluate the conceptual validity of the procedure. Reitan and Wolfson published a series of studies, summarized in Reitan and Wolfson (1999), that showed a definite reduction or even the elimination of significant relationships of age and education in brain-damaged groups, using a summary score that has been shown to be highly sensitive to the effects of brain damage. Among non-brain-damaged groups, however, statistically significant relationships were routinely present. This intergroup difference was presumed to be due to the sensitivity of the summary measure to the various influences of brain damage (severity, duration, lateralization, type of pathology, etc.) and, when brain damage determined the score, the influence of age and education was diminished or overshadowed.

[1]This study was completed before publication of Heaton, R. K., Miller, S. W., Taylor, M. J., and Grant, I. (2004). *Revised comprehensive norms for an expanded Halstead–Reitan Battery: Demographically adjusted neuropsychological norms for African-American and Caucasian adults*. Lutz, FL: Psychological Assessment Resources.

Requests for reprints should be sent to Ralph M. Reitan, Reitan Neuropsychology Laboratory, P.O. Box 66080, Tucson, AZ 85728–6080, USA. E-mail: reitanlabs@aol.com

Reitan and Wolfson (2004) presented a review of publications both supporting and criticizing the HGM transformations. They concluded that methodological problems could influence adversely the application of these transformations, based on a normal group, to persons with brain damage in terms of both psychometric and conceptual considerations.

Despite the widespread use of the HGM transformations in clinical evaluations, research analyses and reports, and legal proceedings, there are few empirical studies to determine the accuracy or limitations of HGM age and education transformations. Golden and van den Brock (1998) conducted a study based on neuropsychological testing of persons with focal brain lesions and concluded that the HGM method tended to produce more normal scores than would be expected based on raw scores. Reitan and Wolfson (2004) also reported that HGM transformations increased the number of normal scores among persons without brain damage and, to an even greater extent, among persons with brain damage. They compared Reitan and Wolfson's Neuropsychological Deficit Scale (RW NDS) scores (Reitan & Wolfson, 1988, 1993) with HGM scores and found significantly more normal scores in both groups when the HGM transformations were used.

RW NDS scores represent categories of performance on Halstead–Reitan Battery variables on a four-point scale (0, 1, 2, and 3), subdivided by a cutting score that corresponds exactly with the cutting score for raw score distributions, that differentiates performances characteristic of persons with and without brain damage. The HGM method converts raw scores into *T* scores, and scores more than a standard deviation below the mean are presumed to be impaired. Thus, an opportunity exists to compare the number of normal and impaired scores that are produced using the two methods. RW NDS scores make no adjustment for age or education but instead represent a perfectly straightforward conversion in each instance, depending only on the value of the raw score. Thus, the number of impaired scores versus normal scores would be exactly the same for RW NDS scores as for raw scores. The HGM procedure uses both age and education, in combination for every individual, in converting the raw score on each test to a *T* score.

The purpose of this study was to convert raw scores on a number of neuropsychological tests, representing performances of a group of persons with definitely established diffuse or bilateral brain disease or damage, into both HGM *T* scores and RW NDS scores and to determine and compare the frequency of normal versus impaired scores when using each method.

METHOD

Participants

A group was composed consisting of 40 men and 12 women, each of whom had a definitive neurological diagnosis of diffuse or bilateral brain disease or damage. Means and standard deviations for age and education were 46.10 years (SD = 12.66) and 11.15 years (SD = 3.89), respectively. The median age was 46.00 years and the median education was 10.50 years.

Diagnoses

Neurological diagnoses were as follows: diffuse cerebral vascular disease, 14; multiple sclerosis, 9; metastatic cerebral tumors, 6; cerebral atrophy, 4; closed-head injury, 4; Alzheimer's disease, 2; bilateral ependymoma, 2; complex partial epilepsy, 2; major motor epilepsy, 1; encephalitis, 1; cavernous hemangioma, 1; dementia paralytica, 1; Wilson's disease, 1; bilateral meningioma, 1; subdural hygroma, 1; third ventricle tumor, 1; and Parkinson's disease, 1. The group was deliberately diversified with respect to diagnoses.

Sixteen of the 52 cases were from the neurosurgical service, and in 10 of these patients the diagnosis was confirmed by surgery and subsequent pathological examination of tissue specimens. In the remaining six neurosurgical cases the lesions or brain abnormality was visualized by brain imaging procedures. Thirty-six of the 52 cases were from the neurology service. All appropriate specialized neurological tests were used, combined with physical neurological examinations, the history, and various additional examinations as needed. Some patients, particularly those with multiple sclerosis, were followed over time to confirm the diagnosis. Only those patients in whom board-certified faculty physicians felt that the diagnoses were unequivocal were included in the study.

Procedure

Archival data were used, based on neuropsychological testing that was done by thoroughly trained technicians whose competence in administration of the tests used was personally verified by the first author. Each technician had been trained in methods for eliciting the best possible performance of each participant and was aware that achievement of this aim was of prime importance in every case. The technicians had no knowledge of the diagnosis of the participants at the time of the testing, or even whether the participant was a patient or a control subject. In addition, the technicians had no knowledge of the research questions for which the test results might be used. Scoring accuracy was verified in an independent review by another technician shortly after testing had been completed. Thus, the data were obtained objectively and were in final form before being used in this study. These procedures were followed to ensure accuracy of the test results and to limit any possible bias in the collection of data for this study.

The procedure involved transformation of raw neuropsychological test scores to HGM *T* scores and RW NDS scores. Next, the frequency of impaired scores versus normal scores was determined for each method for each of the 11 tests included. Although these numbers are presented for each test using each method, the analyses were based on the total number of impaired scores versus the total number of normal scores produced by each method. An additional step involved comparisons of the number of impaired scores versus the number of normal scores in each half of the cases, divided at the median, for age and education. Reitan and Wolfson (2004) had hypothesized that less-educated persons with brain damage and persons of older age would be the principal recipients of normal scores when using the HGM transformations, as supported by data reported by Heaton, Grant, and Matthews (1991). Thus, even though this study was not deliberately designed to assess differential age and education effects, the availability of the data presented this opportunity as an initial evaluation. Inasmuch as the data represented discrete variables (instances of normal versus impaired scores), the χ^2 statistic was used for comparisons.

RESULTS

Table 1 presents mean raw scores and standard deviations for age and education and the 11 variables included in this study for 52 persons with brain damage. The results shown in Table 1 document evidence of significant impairment on all 11 tests, as would be expected on the basis of many prior reported studies of patients with definite brain damage or disease. Thus, as a starting point in this investigation, the raw scores indicated the presence of distinct neuropsychological deficits and would be expected to produce a preponderance of test scores in the impaired range.

Table 2 lists the number and percentage of test scores in the normal range and the impaired range, using conversions based on both the HGM method and the RW NDS method for each of the 11 tests as well as for the total of the 11 tests. Except for the Speech Sounds Perception Test and Part A of the Trail-Making Test, the test results showed conversion frequencies that were consistently greater in the normal range when using the HGM transformations than when using the RW NDS transformations. These results suggest that raw scores that generally reflected impairment (Table 1) were more often converted to normal scores when using the HGM method than when using the RW NDS method. Specifically, the number of normal scores was 1.74 times greater (approaching twice as many normal scores) when using the HGM method than when using the RW NDS method. A χ^2 test of the differential total frequencies yielded a value of 32.66 ($p < .001$). There would seem to be little doubt that the two methods differed in the frequency with which they produced normal scores and impaired scores. Considering the raw score distributions, it appears that a finding of only 64% of the scores falling in the impaired range when using the HGM method may well underestimate the degree of impairment shown by this group of persons with brain damage.

The next comparisons were based on subdivision of the 52 individuals with brain damage into equal-numbered groups of persons with lesser and greater education and younger and older age. Table 3 presents the number and percentage for each of the 11 tests that fell in the normal range and the impaired range for both the HGM conversions and the RW NDS conversions.

The next step in the data analysis was based on comparisons of the lesser educated versus better educated and the younger versus older subgroups. The HGM method produced the same total number of impaired scores and normal scores for the group with lower education and the group with higher education. Similar results were obtained when raw scores were converted for age, with identical numbers of impaired scores and normal scores in both the younger group and the older group.

Thus, when using the HGM transformations for the lesser educated and better educated groups, the results showed the same relationship between impaired scores and normal scores for the subdivided groups and for the total group: 36% of the *T* scores fell in the normal range and 64% fell in the impaired range. Obviously, there was no difference in the frequency of normal scores and impaired scores according to age or education level, and, as compared with the total group, the transformations perfectly achieved the stated aim in adjusting for differing age and education level.

Table 1. *Raw Score Means and Standard Deviations for a Group of 52 Persons With Definite Evidence of Diffuse or Bilateral Cerebral Damage or Disease*

				TPT					Tapping				
	Age	**Education**	**Category**	**Time**	**Memory**	**Loc**	**Rhythm**	**SSPT**	**Dominant**	**Nondominant**	**Impairment Index**	**Trails A**	**Trails B**
M	46.1	11.15	76.48	34.92	3.83	1.35	20.40	17.42	36.73	35.92	0.71	83.79	236.50
SD	12.66	3.89	38.57	18.42	2.99	1.53	6.55	10.99	18.08	12.05	0.22	60.81	176.27

Note. TPT = Tactual Performance Test; SSPT = Speech-sounds Perception Test.

Table 2. *Overall GroupsNumber and Percentage of Scores in the Normal Versus Impaired Range When Transformations Were Made Using HGM T-Scores and RW NDS Scores*

	Category		TPT—Time		TPT—Mem		TPT—Loc		Rhythm		SSPT		Tapping-Dominant		Tapping-Nondominant		Impairment Index		Trails A		Trails B		Total	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Normal HGM T-scores of 40+	27	52	21	40	15	29	17	33	23	44	13	25	22	42	29	56	14	27	11	21	13	25	205	36
Impaired HGM T-scores of 39–	25	48	31	60	37	71	35	67	29	56	39	75	30	58	23	44	38	73	41	79	39	75	367	64
Normal RW NDS scores of 0 or 1	16	31	7	13	10	19	1	2	19	37	17	33	10	19	12	23	8	15	11	21	7	13	118	21
Impaired RW NDS scores of 2 or 3	36	69	45	87	42	81	51	98	33	63	35	67	42	81	40	77	44	85	41	79	45	87	454	79

Note. HGM = Heaton, Grant, and Matthews; RW NDS = Reitan and Wolfson's Neuropsychological Deficit Scale; TPT = Tactual Performance Test; SSPT = Speech-sounds Perception Test.

Table 3. *Number and Percentage of HGM T Scores in the Normal and Impaired Range According to Education and Age Categories*

	Category		TPT—Time		TPT—Mem		TPT—Loc		Rhythm		SSPT		Tapping-Dominant		Tapping-Nondominant		Impairment Index		Trails A		Trails B		Total	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Lesser Education																								
HGM *T* scores of 40+	13	50	12	46	11	42	10	38	9	35	5	19	9	35	14	54	8	31	4	15	8	31	103	36
HGM *T* scores of 39–	13	50	14	54	15	58	16	62	17	65	21	81	17	65	12	46	18	69	22	85	18	69	183	64
Greater Education																								
HGM *T* scores of 40+	11	42	10	38	4	15	7	27	14	54	8	31	12	46	16	62	8	31	7	27	6	23	103	36
HGM *T* scores of 39–	15	58	16	62	22	85	19	73	12	46	18	69	14	54	10	38	18	69	19	73	20	77	183	64
Lesser Age																								
HGM *T* scores of 40+	15	58	11	42	7	27	8	31	12	46	5	19	9	35	14	54	5	19	5	19	7	27	98	34
HGM *T* scores of 39–	11	42	15	58	19	73	18	69	14	54	21	81	17	65	12	46	21	81	21	81	19	73	188	66
Greater Age																								
HGM *T* scores of 40+	10	38	10	38	8	31	10	38	11	42	8	31	12	46	10	38	9	35	6	23	4	15	98	34
HGM *T* scores of 39–	16	62	16	62	18	69	16	62	15	58	18	69	14	54	16	62	17	65	20	77	22	85	188	66
Lesser Education																								
RW NDS Scores–Normal	6	23	5	19	5	19	1	4	5	19	4	15	4	15	4	15	3	12	3	12	3	12	43	15
RW NDS Scores– Impaired	20	77	21	81	21	81	25	96	21	81	22	85	22	85	22	85	23	88	23	88	23	88	243	85
Greater Education																								
RW NDS Scores–Normal	10	38	2	8	4	15	0	0	12	46	13	50	6	23	8	31	5	19	8	31	4	15	72	25
RW NDS Scores–Impaired	16	62	24	92	22	85	26	100	14	54	13	50	20	77	18	69	21	81	18	69	22	85	214	75
Lesser Age																								
RW NDS Scores–Normal	10	38	6	23	7	27	0	0	13	50	10	38	7	27	6	23	5	19	8	31	7	27	79	28
RW NDS Scores–Impaired	16	62	20	77	19	73	26	100	13	50	16	62	19	73	20	77	21	81	18	69	19	73	207	72
Greater Age																								
RW NDS Scores–Normal	6	23	1	4	2	15	1	4	6	23	7	27	4	15	6	23	2	15	3	12	0	0	38	13
RW NDS Scores–Impaired	20	77	25	96	24	85	25	96	20	77	19	73	22	85	20	77	24	85	23	88	26	100	248	87

Note. HGM = Heaton, Grant, and Matthews; RW NDS = Reitan and Wolfson's Neuropsychological Deficit Scale; TPT = Tactual Performance Test; SSPT = Speech-sounds Perception Test.

The RW NDS score transformations for these same groups classified only 15% of the scores in the normal range for the group with less education and 25% in the normal range for the group with more education. This difference yielded a χ^2 value of 9.15 ($p < .005$), indicating that the group with more education had significantly more normal scores than impaired scores as compared with the group with less education, even though both groups had fewer normal scores than were produced with the HGM transformations.

When the number of impaired scores and normal scores for the younger group and older group was determined on the basis of RW NDS scores, the younger group earned impaired scores in 72% of the instances and normal scores in 28% of the instances. The older group earned impaired scores in 87% of the instances and normal scores in 13%. These differences in the younger and older groups yielded a χ^2 value of 18.08 ($p < .001$), indicating that the RW NDS conversions produced impaired scores more frequently in the older group than in the younger group at a statistically significant level.

Thus, whereas the HGM transformations resulted in exactly the same number of impaired scores and normal scores for the subgroups divided by age or education, the RW NDS transformations resulted in significantly more persons showing impaired scores in the subgroups with lesser education and older age.

The preceding results evaluated differences relating to age and education within groups, but comparisons of perhaps greater interest involve differences between HGM and RW NDS methods. Table 4 presents comparisons of the number of impaired scores (stated in percentages) produced by the two methods according to age and education differences. Significant differences in the number of impaired scores were found in three of the four comparisons. It is noteworthy that the greatest differences between the two methods occurred in the subgroups with lower education and older age. In each of these subgroups the RW NDS method produced substantially more impaired scores than did the HGM method, yielding χ^2 values of 33.11 and 34.73. However, in every subgroup the RW NDS method produced a higher percentage of impaired scores than did the HGM method, although the difference was not significant for the younger groups.

DISCUSSION

First, it is apparent that the HGM *T* score transformations produced significantly more normal scores in our brain-damaged sample than did the RW NDS transformations. This finding clearly implies that the neuropsychological correlates of brain damage were identified less often when using the HGM method, suggesting a lesser sensitivity than when using the RW NDS method. No conclusions can be drawn from this study about specificity or about the accuracy of the two methods in differentiating brain-damaged participants from non-brain-damaged controls, inasmuch as the latter group was not included in this study. However, the findings do have definite clinical and forensic implications. The results of this study suggest that using the HGM method rather than the RW NDS method is more likely to produce conclusions of normality rather than impairment among persons with diffuse or bilateral brain damage. A host of studies over the years have shown that impairment rather than normality is to be expected among persons with brain damage; however, many individuals with brain damage do, in fact, produce some normal scores. Thus, many additional questions need to be investigated to understand more fully the consequences of various methods of raw score transformations. Until these questions are answered more completely, it would seem wise to rely on methods of interpretation, based on raw scores, that have been repeatedly validated and for which sensitivity and specificity results have been reported for differentiating groups with and without brain damage (Reitan & Wolfson, 1988, 1993). In addition, research publications based on HGM transformations using the tests studied in this investigation may very well be based on data that underestimate the impairment experienced by persons with brain damage. The likelihood of

Table 4. *χ^2 Comparisons of Age and Education Effects on Production of Impaired Scores When Using the HGM Method as Compared With the RW Method of Transforming Raw Scores*

	χ^2	$p <$
Lower education: HGM (64% impaired) vs. RW NDS (85% impaired)	33.11	.001
Higher education: HGM (64% impaired) vs. RW NDS (75% impaired)	7.91	.005
Younger age: HGM (66% impaired) vs. RW NDS (72% impaired)	2.95	.10
Older age: HGM (66% impaired) vs. RW NDS (87% impaired)	34.73	.001

Note. RW = Reitan–Wolfson; HGM = Heaton, Grant, and Matthews.

underestimation of the neuropsychological effects of brain damage in cases involved in litigation, when using the HGM transformations, is in even more urgent need of clarification.

One of the postulates of Reitan and Wolfson (2004), based on psychometric and conceptual considerations, was that the corrections produced by the HGM method would be reflected principally among persons with lesser education and greater age. This postulate was confirmed by the findings of this study through comparisons of HGM transformations with scores produced by the RW NDS method (which directly reflects the actual raw scores).

A likely reason for the postulate is that older age and lesser education both are variables that could be associated with brain impairment and resulting neuropsychological impairment. With respect to age, Reitan and Wolfson (1986) have reviewed many studies of persons who had no medical evidence of brain damage but nevertheless showed definite age-related evidence of neuropsychological impairment. Neuropathological findings on autopsy of presumably normal persons have shown many changes of the type seen in Alzheimer's disease, not only among individuals in North America and Europe, 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 individuals 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. Although the brains of persons with lower education who are presumed to be normal neurologically have not, to our knowledge, been studied in as much detail as the brains of older persons, it is reasonable to presume that lesser educational achievement, among other factors, may be limited by lesser competence of brain functions and certainly would be associated with lower scores on neuropsychological tests.

Inasmuch as the HGM method brings about its greatest corrections among persons whose scores are more likely to be affected by brain impairment, a question must be raised concerning exactly what the HGM transformations are accomplishing. It might seem, at least in part, that the corrections are, in fact, correcting for subtle impairment of brain functions in less-educated and older persons—the very condition that neuropsychological tests were developed to detect.

The questions in this area are obviously complex and multifaceted. This study included only persons with diffuse or bilateral cerebral involvement. Similar studies need to be done with older and younger normal persons and, considering the many dimensions of brain damage and disease, with many groups that differ in these respects. The ultimate interest of neuropsychologists is the correct and valid evaluation of persons with brain disease, damage, or dysfunction, and investigations based on persons with such conditions are especially needed.

Finally, it should be noted explicitly that nothing in this study should be construed as opposing the perfectly valid procedure of comparing the results of any individual (or group) with data derived from a similar or different category (age, education, gender, diagnosis, etc.). Our concern was not with such considerations. Instead, the purpose of this study was directed toward the concurrence, or level of agreement, between raw-score transformations when generated by methods that differed in producing such transformations.

REFERENCES

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., & Matthews, C. G. (1991). *Comprehensive norms for an expanded Halstead–Reitan Battery.* Odessa, FL: Psychological Assessment Resources.

Jarvis, P .E., & Barth, J. T. 1993. *The Halstead–Reitan Neuropsychological Test 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.

Reitan, R. M., & Wolfson, D. (1986). 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. (1988). *Traumatic brain injury. Vol. II. Recovery and rehabilitation.* Tucson, AZ: Neuropsychology Press.

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

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.

Reitan, R. M., & Wolfson, D. (2004). Clinical and forensic issues regarding age, education, and the validity of neuropsychological test results: A review and presentation of a new study. *Journal of Forensic Neuropsychology, 4,* 1–32.