Reference data for the Word Memory Test

Rienstra, A.; Spaan, P.E.J.; Schmand, B.A.

Published in:
Archives of Clinical Neuropsychology

DOI:
10.1093/arclin/acp035

OXFORD JOURNALS
OXFORD UNIVERSITY PRESS

Archives of Clinical Neuropsychology 24 (2009) 255–262

Archives of CLINICAL NEUROPSYCHOLOGY

Reference Data for the Word Memory Test

A. Rienstra[a,b,*], P. E. J. Spaan[b], B. Schmand[a,b]

[a]*Department of Neurology, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands*
[b]*Department of Psychonomics, University of Amsterdam, Amsterdam, the Netherlands*

Accepted 18 June 2009

Abstract

Many studies have evaluated the utility of the Word Memory Test (WMT) as a symptom validity test. However, there is a lack of reference data for the WMT conventional memory subtests. The present study examined the demographic characteristics that influence performance on these subtests, in order to develop demographically corrected reference data. For this purpose, we administered the Dutch version of the WMT to 115 healthy Dutch controls, aged 20–80 years. Furthermore, we demonstrated the equivalence of the English and Dutch language versions of the WMT. Stepwise linear regression analyses of the combined Canadian and Dutch samples ($N = 155$) showed that the memory scores declined with increasing age. Participants with lower levels of education performed worse than more highly educated subjects. Reference data stratified by age and level of education are presented for use in research and clinical settings.

Keywords: Word Memory Test; Reference data; Memory; Neuropsychological assessment; Symptom validity testing

Introduction

Memory complaints are the most frequent reason for neuropsychological referral in outpatient settings (Lezak, Howieson, & Loring, 2004). The assessment of memory deficits, however, can be seriously complicated by suboptimal performance. Numerous studies have shown substantial exaggeration of cognitive impairment in certain groups of people being given neuropsychological tests, especially when financial claims are involved (Larrabee, 2000). Green, Rohling, Allen, and Lees-Haley (2001) stated that approximately 50% of variance of performance on a neuropsychological battery administered to a heterogeneous group of disability claimants is explained by effort and cooperation. Subsequent studies in the USA and Germany independently replicated this finding (Constantinou, Bauer, Ashendorf, Fisher, & McCaffrey, 2005; Stevens, Friedel, Mehren, & Merten, 2008), confirming that the effect of effort on neuropsychological test scores is more pronounced than the effect of brain injury. Therefore, it is recommended that objective tests of effort and associated symptom validity are used in neuropsychological assessment, especially when potential secondary gain increases the incentive for symptom exaggeration as well as when neuropsychologists are suspicious of insufficient effort, or inaccurate or incomplete responding (Bush et al., 2005).

According to Hartman (2002), the Word Memory Test (WMT) is one of the most popular and best investigated symptom validity tests currently available (Green, 2003; Green, Allen, & Astner, 1996). The WMT is a word list-learning task containing multiple subtests, of which the first two are specifically designed to evaluate effort, while the remaining subtests are conventional tests of verbal memory (see Materials and Methods section for details). The WMT effort subtests have been validated in a large number of studies (see www.wordmemorytest.com for an overview of the literature). For example, it has been shown consistently that many patients with brain injury and brain diseases score well above the cut-offs for incomplete effort (Allen & Green, 1999; Green & Allen, 1999; Green, Iverson, & Allen, 1999; Iverson, Green, & Gervais, 1999). In addition,

* Corresponding author at: Department of Neurology H2-222, Academic Medical Center, P.O. Box 22660, 1100 DD Amsterdam, the Netherlands. Tel.: +31-20-56-66822; fax: +31-20-56-69217.
E-mail address: a.rienstra@amc.uva.nl (A. Rienstra).

doi:10.1093/arclin/acp035 Advance Access publication on 18 July 2009

there were no differences on the WMT effort measures between neurological patients with and without impaired verbal memory (Green et al., 1999). Finally, in several independent studies, the WMT was 99%–100% accurate in discriminating between persons asked to make a good effort and those who were instructed to feign believable memory impairment (Brockhaus & Merten, 2004; Iverson et al., 1999; Jing, Slick, Strauss, & Hultsch, 2002).

Whereas many studies have investigated the utility of the WMT as a symptom validity test, there is a lack of research presenting reference data for the conventional memory subtests (Strauss, Sherman, & Spreen, 2006). Availability of appropriate reference values for these measures would have two important advantages for the use of the WMT in clinical practice. First, improved reference data would promote the efficiency of neuropsychological evaluation because both effort quality and verbal memory can be investigated with a single test. Second, better reference tables will lead to improved interpretation of the WMT profile of subtest scores. Green (2003) states that the shape of the profile of WMT subtest scores (in particular the difference between the effort subtests and conventional memory subtests) indicates whether patients fail due to insufficient effort or to the severity of their cognitive disorders. Therefore, appropriate reference data are of crucial importance to judge whether or not an individual's profile is consistent with his age and education.

The WMT user's manual (Green, 2003) provides mean scores of different comparison groups, including normal adult controls, but the sample size of this group is rather small and has a limited age range. Furthermore, although the WMT has been translated into several languages, the normal adult control group was composed of native English speakers who were given the original English WMT. One cannot assume that different language versions of the test are of equal difficulty. This should be established empirically. In addition, the relation between the WMT memory subtests and demographic variables has not been studied thoroughly. Like many other cognitive tests, the conventional WMT subtests are probably influenced by age, education and, perhaps, gender. Indeed, significant effects of age and educational level were found for all WMT memory subtests in the control sample, but the influence of gender was not investigated and there were only a few healthy subjects over 60 years of age (Green, 2003).

The purpose of the present study was twofold. Our first aim was to investigate the equivalence of the English and Dutch versions of the WMT by comparing the data of the original Canadian normative sample with those of a matched comparison group of Dutch adult normal controls. Second, we examined the demographic characteristics (age, gender, and level of education) that may influence performance, in order to create a demographically corrected data set. In line with previous research (Green et al., 1996; Green, Lees-Haley, & Allen, 2002) we expected that the scores on the effort subtests are unrelated to demographic variables. On the other hand, consistent with the literature on tests of verbal memory such as Rey's Auditory Verbal Learning Test (Messinis, Tsakona, Malefaki, & Papathanasopoulos, 2007; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2005; Schmidt, 1996) we hypothesized that performance on the conventional memory measures does correlate with demographic factors. Our hypotheses were that performance on all memory measures is (a) negatively related to age; (b) positively to the level of education; and (c) is associated with gender in such a way that women perform better than men.

Materials and Methods

Participants

Two groups of volunteer participants were involved in this study. The first group consisted of 117 healthy Dutch adults, who took part in the study after providing written consent. To obtain a nationally representative sample of the Dutch population, the research sample was stratified for gender, three age levels, and three levels of education based on data from Statistics Netherlands (Centraal Bureau voor de Statistiek, 2006; http://statline.cbs.nl). The cell frequencies were made to correspond to the percentages of the Dutch population in the respective cells. Appropriate subjects were recruited in the circle of acquaintances of the investigators and subsequently the snow ball method was used, that is each participant was asked to provide names of other potential subjects. In line with the data from Statistics Netherlands, age groups ranged from 20 to 40, 41 to 65, and 66 to 80 years with corresponding average ages of 28.6 ($SD = 7.1$), 55.3 ($SD = 7.6$), and 72.9 ($SD = 4.2$), respectively. The low level of education group (LE low) consisted of persons with primary education only, the average level of education group (LE average) represented junior vocational training, and the high level of education group (LE high) contained participants with senior vocational or academic training. These three levels of education corresponded in our sample with an average of, respectively, 9.5 ($SD = 0.9$), 13.5 ($SD = 0.9$), and 17.7 ($SD = 1.5$) years of full-time formal education. Exclusion criteria were: (a) a history of physical or psychiatric disorders that could interfere with cognitive performance including neurological disorders or history of concussion; (b) pre-existent mental retardation; (c) the chronic use of psychotropic drugs; (d) substance abuse or alcohol intake of more than three units per day; (e) physical disability that interferes with neuropsychological testing, and (f) Dutch as a second language.

The second group consisted of 40 healthy English-speaking Canadian adult volunteers (Iverson et al., 1999). They were recruited by an advertisement posted on notice boards in office buildings and a newspaper. They were offered a free $10 gift token for a local restaurant. If they took the tests, they got the free lunch. All participants were healthy functioning adults and fluent in English. Exclusion was based on self-report of no psychiatric history and no neurological diseases.

Measures

The computerized assessment of the WMT requires participants to memorize a list of 20 simple semantically related word pairs (e.g., "man" and "woman"). After the list is presented twice, a forced-choice Immediate Recognition (IR) subtest is displayed in which the person is asked to select the word from the original list (e.g., "man" from the word pair "man–boy"). Without advance warning, after a 30-min delay, the forced-choice Delayed Recognition (DR) subtest is administered in the same way as the IR, except that different foil words are used (e.g., "man–teenager"). On the basis of the performance on the IR and DR subtests, the computer calculates the consistency score (CNS) as percent agreement in responses between these two trials. Any score on the IR, DR, and CNS measures below the recommended cut-off point is indicative of suboptimal effort. The effort measures are followed by four subtests of gradually increasing difficulty, which measure verbal memory ability. These are conventional memory tests. The first is the Multiple Choice (MC) subtest in which the first word from each word pair is shown and the correct second word has to be chosen from eight options. On the Paired Associates (PA) subtest, the first word from each word pair is given by the investigator and the subject is asked to say the second word. The last two subtests are the Free Recall (FR) and the Long Delayed Free Recall (LDFR) subtests, in which participants are asked to recall as many words as possible from the original list of word pairs. The FR subtest is given immediately after the PA subtest and LDFR after another 20-min interval. Auditory and visual feedbacks are given after each response on the IR, DR, and MC subtests. These measures, therefore, serve not only as effort subtests but also as additional learning trials, which assists motivated participants in learning. The PA subtest provides further exposure to the first words of each pair, again serving a dual purpose as a test of memory and as a learning trial, prior to FR.

Procedure

All Dutch participants were tested individually at their homes by a psychology student of the University of Amsterdam. Students taking part in the testing of participants had been intensively trained in the administration procedures of the psychological measures involved (see subsequently), and they were supervised by the first author (A.R.). Subjects were screened by means of a standardized interview at the beginning of the testing session, in order to exclude those with health problems or other exclusion criteria as described earlier. Participants were assessed using the Dutch translation of the computerized version of the WMT (Green, 2003). Test instructions were given on screen, according to the guidelines for test administration in the WMT user's manual. In the 30-min period between the IR and the DR trials and in the 20-min interval between the FR and the LDFR subtests, other non-verbal neuropsychological tests were administered to avoid interference with the WMT material. This battery included the Trail Making Test (Reitan, 1992), the Stroop Color Word test (Stroop, 1935), the Visual Reproduction subtest of the Wechsler Memory Scale-Third edition (Wechsler, 1997), and the Non-verbal Medical Symptom Validity test (Green, 2004). The complete assessment took approximately an hour and a half.

The Canadian participants were tested by a trained psychological assistant at a professional psychology office, applying the standard administration. An intelligence subtest was administered in the delay period after the IR subtest.

Statistical Analysis

Participants who failed the effort subtests of the WMT were excluded from the analysis. Incomplete effort was defined by the cut-off points recommended in the WMT user's manual (Green, 2003). To investigate the equality of means, paired samples *t*-tests were used for normally distributed variables. The Pearson χ^2 test and the Wilcoxon Signed Ranks test for paired data were employed for variables that were not normally distributed. Bonferroni correction of the level of significance was used to correct for multiple comparisons. Pearson's r or Spearman's rho were used to calculate correlation coefficients. To asses possible relationships between the conventional memory subtests of the WMT and demographic variables (age, gender, and years of education), stepwise multiple linear regression models were fitted, with the memory subtests of the WMT as dependent and the demographic variables as independent variables. Van der Waerden-normalized scores were used when the data were skewed and the transformed scores were applied in the regression analyses. All analyses were conducted using SPSS (version 14.0) for Windows. For all tests, values less than $.05/6 = .008$ were accepted as significant.

Results

In the Dutch sample of 117 healthy adults, 2 participants failed the WMT based on the recommended cut-offs. These subjects were excluded, because it is very likely that their WMT profiles were invalid. The first case was a 40-year-old man with 14 years of education. He scored lower than an early dementia group on the easy subtests, while his scores on the more difficult subtests (FR) were within the normal range. This represents internal inconsistency, a function of inconsistent effort. The second case was a woman, 79 years of age, who has had 14 years of education. She produced a very irregular profile that could be typical for dementia.

In the Canadian group, all 40 cases passed the three WMT effort measures. As a result, a total of 115 Dutch participants and 40 Canadian subjects were included in this study. The Dutch sample was 47% men ($N = 54$), average age 49.8 years ($SD = 18.4$, range: 20–80), and had 13.9 years of education ($SD = 3.3$, range: 7–19). As can be seen in Table 1, their scores on other neuropsychological tests administered were in the normal range. Forty-five percent of the Canadian sample was men ($N = 18$), the average age was 36.8 years ($SD = 11.4$, range: 16–68), and they had an average of 14.6 years of formal education ($SD = 1.7$, range: 10–16). Their mean vocabulary subtest scaled score on the WAIS-R was 11.8 (SD = 3.9).

Equivalent Difficulty of the Dutch and Canadian WMT Subtest Means

To compare the mean WMT subtests scores of the Dutch and the Canadian participants, each Canadian was matched to a Dutchman by gender, age, and years of education. Five Canadian subjects could not be matched because they were younger than 20 years of age. This resulted in a total of 35 subjects in both groups. After the matching procedure, there were no significant differences between groups (gender: χ^2 (df = 1) = .21, $p = .65$; age: $Z = -1.22$, $p = .22$; education: $Z = -0.99$, $p = .32$).

Fig. 1 shows the mean WMT subtest scores of the matched Dutch and Canadian sample.

We found no significant differences between groups on any of the WMT subtests (IR: $Z = -2.56$, $p = .011$; DR: $Z = -0.37$, $p = .71$; CNS: $t(39) = 1.33$, $p = .19$; MC: $Z = -0.53$, $p = .59$; PA: $Z = -0.92$, $p = .36$; FR: $t(34) = 1.65$, $p = .11$; LDFR: $t(34) = 0.35$, $p = .73$).

Table 1. Standard cognitive test scores for the Dutch sample ($n = 117$)

Stroop test word reading	53.9 (11.6)[a]	(28.7–80.9)
Stroop test color naming	54.9 (11.7)[a]	(30.1–86.9)
Stroop interference test	54.8 (12.4)[a]	(15.6–85.9)
Trail Making Test A	52.9 (11.2)[a]	(25.8–82.3)
Trail Making Test B	55.3 (9.9)[a]	(34.0–76.3)
WMS-III VR immediate	10.0 (3.1)[b]	(2–18)
WMS-III VR delayed	11.6 (2.7)[b]	(5–18)
WMS-III percent retained	12.2 (2.6)[b]	(5–18)

Notes: Values are mean (*SD*) and range.
WMS-III VR = Wechsler Memory Scale–Third edition Visual Reproduction subtest.
[a]T-scores corrected for age, education and gender.
[b]Wechsler Scaled-Scores corrected for age.

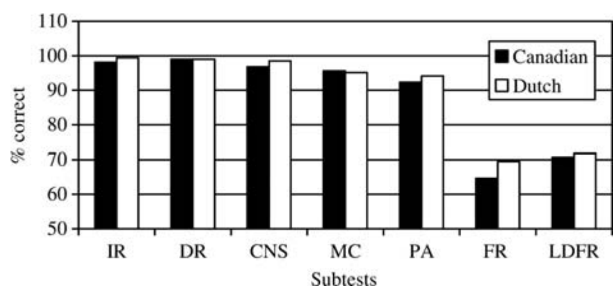

Fig. 1. Mean WMT subtest scores of the matched Canadian and Dutch samples.

Influence of Demographic Variables on WMT Subtests

Given the high similarity between the mean WMT subtest scores of the matched Dutch and Canadian subjects, we decided to combine both data sets. This resulted in a total sample of 155 participants. Descriptive data for the entire sample are provided in Table 2.

Table 3 shows the correlations between gender, age, and years of formal education and the WMT subtests.

No significant associations were found between any of the WMT subtest scores and gender. However, the scores on the WMT effort measures as well as the performance on the conventional memory subtests of the WMT were significantly related to age and years of formal education.

A series of stepwise multiple regression analyses was conducted with each of the memory subtests of the WMT as the dependent variable, and gender, age, and years of education as the predictors. The scores on the MC and PA subtests were Van der Waerden transformed into z-scores. Table 4 presents the regression weights, the standard errors, and the proportions of explained variance for the entire sample.

The pattern of beta weights was comparable with the correlations in Table 3, with age and years of education both making significant contributions to the prediction of all WMT memory scores. In general, performance on the WMT memory subtests declined with increasing age and more highly educated participants scored significantly better than less-educated subjects. The R^2 values indicated that age and education explained up to 20% of test score variance. Gender only made a significant contribution to the prediction of the LDFR subtest score, explaining together with age and years of education 24% of test

Table 2. Descriptive characteristics of the sample ($N = 155$)

Age group	Age			Level of education			Male:female ratio
	N	*M*	*SD*	Low	Average	High	
20–40	65	29.03	7.47	12	28	25	31:34
41–65	63	53.27	7.83	10	28	25	27:36
66–80	27	72.44	4.11	6	13	8	14:13
Total	155	46.45	17.78	28	69	58	72:83

Table 3. Correlation coefficients gender, age, and education and WMT subtests

	IR[a]	DR[a]	CNS[a]	MC[a]	PA[a]	FR[b]	LDFR[b]
Gender	−.01	.02	.01	.03	−.04	.11	.12
Age	−.22[c]	−.18[d]	−.27[c]	−.34[c]	−.32[c]	−.36[c]	−.38[c]
Education	.24[c]	.16[d]	.31[c]	.32[c]	.30[c]	.27[c]	.29[c]

Notes: [a]Spearman rank correlation.
[b]Pearson's *r* correlation.
[c]Correlation is significant at the .01 level (two-tailed).
[d]Correlation is significant at the .05 level (two-tailed).

Table 4. Multiple linear regression models of the WMT memory subtests following a step-down procedure

	Score	Variable	*B*	Std. error *B*	*t*	*p*	R^2	*SEst*
Memory subtests WMT	MC	(Constant)	.346	.109	3.165	.002	.202	.246
		Age	−.005	.001	−4.838	<.001		
		Education	.029	.007	4.273	<.001		
	PA	(Constant)	.330	.111	2.984	.003	.191	.249
		Age	−.005	.001	−4.551	<.001		
		Education	.029	.007	4.263	<.001		
	FR	(Constant)	59.358	6.079	9.765	<.001	.193	13.72
		Age	−.313	.062	−5.023	<.001		
		Education	1.400	.373	3.754	<.001		
	LDFR	(Constant)	51.194	7.644	6.697	<.001	.239	14.89
		Age	−.373	.068	−5.517	<.001		
		Education	1.756	.405	4.335	<.001		
		Gender	4.777	2.402	1.988	.049		

Table 5. Means and *SD*'s for WMT subtests stratified by age and level of education

Level of education	Age	IR	DR	CNS	MC	PA	FR	LDFR
Low	20–40	98.4 (2.2)	98.4 (2.3)	97.2 (3.6)	91.3 (6.4)	88.3 (9.4)	65.3 (9.4)	62.8 (22.5)
	41–65	97.1 (4.1)	98.6 (3.2)	96.7 (4.2)	88.5 (11.6)	82.0 (17.2)	65.6 (17.7)	60.0 (15.9)
	66–80	97.5 (2.2)	97.2 (2.5)	95.0 (3.5)	80.0 (12.3)	78.3 (12.1)	50.9 (8.5)	53.1 (9.7)
Average	20–40	98.3 (2.7)	98.4 (2.9)	98.2 (3.2)	94.8 (7.3)	92.5 (10.5)	65.5 (16.4)	67.3 (16.9)
	41–65	98.0 (2.9)	97.8 (2.8)	95.5 (4.4)	89.1 (8.3)	87.1 (10.4)	56.8 (12.5)	59.5 (16.1)
	66–80	96.3 (4.2)	98.0 (2.9)	95.3 (4.7)	87.3 (13.6)	85.0 (14.4)	57.6 (15.9)	57.7 (15.2)
High	20–40	99.7 (.82)	99.3 (1.3)	98.8 (1.9)	97.6 (3.5)	97.2 (3.3)	78.4 (12.2)	81.6 (11.8)
	41–65	98.8 (2.1)	99.2 (1.4)	99.7 (.85)	96.6 (5.5)	94.8 (6.2)	68.2 (11.2)	71.5 (9.0)
	66–80	96.6 (4.4)	96.1 (4.7)	95.8 (5.1)	82.5 (16.7)	81.3 (20.1)	54.5 (13.9)	55.3 (18.5)

score variance. Table 5 provides means and standard deviations for all WMT subtests stratified by age and level of education. Regression equations to calculate standard scores for the FR and LDFR subtests are given in the Appendix.

Discussion

This study provides improved reference data for the WMT memory measures. We extended the sample of normal adult controls and investigated the influence of demographic characteristics. As hypothesized, age as well as education accounted for a substantial proportion of the variance in performance on every memory subtest. More specifically, memory scores declined with increasing age, and participants with higher levels of education performed better than participants with lower educational levels. These findings are consistent with the reports in the literature that age and years of formal education influence performance on tests of verbal memory, such as Rey's Auditory Verbal Learning Test (Messinis et al., 2007; Schmidt, 1996; Van der Elst et al., 2005). Contrary to expectations, gender only explained a small proportion of the variance in performance on the LDFR subtest.

Previous research showed that performance on the WMT effort subtests is almost entirely independent of age and education (Green et al., 1996, 2002). Yet, the current results did show that these subtests were significantly related to these variables. A possible explanation for this discrepancy may be the broader age range of the Dutch sample (20–80 years), in comparison with previously used research groups. In younger age groups, effects of demographic variables might be masked by ceiling effects. In contrast, elderly participants show decline of performance with advancing age. Also, it may be noted that, although statistically significant, the variations in scores on the IR and DR subtests by age and education are clinically insignificant. In Table 5, it can be seen that the mean IR and DR values vary between 96% and 99.5% in all cases, which is a range of less than four percentage points. In contrast, on the FR and LDFR subtests, the scores varied between 50% and 82% correct, representing a range of 32 percentage points. Consequently, even though we picked up some effect of age and education on IR and DR, this does not have implications for the use of the WMT effort measures in clinical practice, because the lowest scores were still above the cut-off score for insufficient effort.

In evaluating the generalizability of the present findings, several potential limitations need to be mentioned. One limitation concerns the question whether it was justified to combine the Canadian and the Dutch data sets. The differences in mean subtest scores between the two language versions were not significant after matching the Dutch and Canadians for age, gender, and years of education. Without the Bonferroni correction, however, the Dutch sample scored slightly, but significantly, higher than the Canadian participants on the IR subtest. Nevertheless, it still is justified to combine the two data sets. The difference in mean IR score is very small, and it is actually irrelevant, because IR is an effort measure where a score above the cut-off is the only thing that matters. Another limitation is the small sample size in some of the cells of the age by education matrix. It would have been desirable to have included more subjects in general and especially more elderly participants with a low level of education. Finally, it is questioned whether the reference data of the English and Dutch forms are applicable to other language versions of the WMT. Future research should establish this empirically.

In conclusion, our study showed significant effects of age and years of education on WMT conventional memory subtests. The results suggest that these two variables need to be taken into account, to evaluate an individual's memory performance accurately. Furthermore, we expanded the sample of adult normal controls and presented reference data stratified by age and level of education. As mentioned in the Introduction section, this has enlarged the area of application of the WMT. These reference data may facilitate a more efficient use of the conventional memory subtests. However, more reference data are welcomed, because the size of the present sample still is too small for a proper standardization.

Funding

Supported by Psychology Research Institute University of Amsterdam.

Conflict of Interest

None declared.

Acknowledgements

The authors thank Dr. P. Green and Dr. R. Gervais for providing the original raw data from normal controls. They also thank Z. van Burkom, T. Erdmann, C. Gomes, T. Kardoes, and R. Schouw for neuropsychological testing of the Dutch control subjects.

Appendix

(1) Standard scores FR subtest

$$\text{Expected FR score} = 59.4 - 0.31 \times \text{age} + 1.4 \times \text{education}.$$

$$\text{T-score FR} = 50 + 10 \times \frac{\text{observed FR} - \text{expected FR}}{\text{SEE}}$$

$$\text{T-score FR} = 50 + 10 \times \frac{\text{observed FR} - (59.4 - 0.31 \times \text{age} + 1.4 \times \text{education})}{13.7}$$

(2) Standard scores LDFR

$$\text{Expected LDFR score} = 51.2 - 0.37 \times \text{age} + 1.76 \times \text{education} + 4.78 \times \text{gender}$$

$$\text{T-score LDFR} = 50 + 10 \times \frac{\text{observed LDFR} - \text{expected LDFR}}{\text{SEE}}$$

$$\text{T-score LDFR} = 50 + 10 \times \frac{\text{observed LDFR} - (51.2 - 0.37 \times \text{age} + 1.76 \times \text{education} + 4.78 \times \text{gender})}{14.9}$$

N.B. Education = years of full-time formal education.

References

Allen, L. M., & Green, P. (1999). Severe TBI sample performance on CARB and the WMT. In P. Green, & L. M. Allen (Eds.), *1999 Supplement for the Computerized Assessment of Response Bias, Word Memory Test and Memory Complaints Inventory.* Durham, NC: CogniSyst, Inc.

Brockhaus, R., & Merten, T. (2004). Neuropsychologische Diagnostik suboptimalen Leistungsverhaltens mit dem Word Memory Test. *Der Nervenarzt*, *75*, 882–887.

Bush, S. S., Ruff, R. M., Troster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H. *et al.* (2005). Symptom validity assessment: Practice issues and medical necessity: NAN Policy & Planning Committee. *Archives of Clinical Neuropsychology*, *20*, 419–426.

Centraal Bureau voor de Statistiek. (2006). Statistical Yearbook 2006. The Hague: Statistics Netherlands.

Constantinou, M., Bauer, L., Ashendorf, L., Fisher, J. M., & McCaffrey, R. J. (2005). Is poor performance on recognition memory effort measures indicative of generalized poor performance on neuropsychological tests? *Archives of Clinical Neuropsychology*, *20*, 191–198.

Green, P. (2003). *Green's Word Memory Test for Microsoft Windows: User's manual.* Edmonton, Canada: Green's Publishing Inc.

Green, P. (2004). *Green's Medical Symptom Validity Test (MSVT) for Microsoft Windows: User's manual.* Edmonton, Canada: Green's Publishing Inc.

Green, P., & Allen, L. M. (1999). Performance of neurological patients on the Word Memory Test (WMT) and the Computerized Assessment of Response Bias (CARB). In P. Green, & L. M. Allen (Eds.), *1999 Supplement for the Computerized Assessment of Response Bias, Word Memory Test and Memory Complaints Inventory.* Durham, NC: CogniSyst, Inc.

Green, P., Allen, L. M., & Astner, K. (1996). *The Word Memory Test: A user's guide to the oral and computer-administered forms, US Version 1.1.* Durham, NC: CogniSyst, Inc.

Green, P., Iverson, G. L., & Allen, L. (1999). Detecting malingering in head injury litigation with the Word Memory Test. *Brain Injury*, *13*, 813–819.

Green, P., Lees-Haley, P. R., & Allen, L. M. (2002). The Word Memory Test and the Validity of Neuropsychological Test Scores. *Forensic Neuropsychology*, *2*, 97–124.

Green, P., Rohling, M. L., Lees-Haley, P. R., & Allen, L. M. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury*, *15*, 1045–1060.

Hartman, D. E. (2002). The unexamined lie is a lie worth fibbing Neuropsychological malingering and the Word Memory Test. *Archives of Clinical Neuropsychology*, *17*, 709–714.

Iverson, G., Green, P., & Gervais, R. (1999). Using the Word Memory Test to detect biased responding in head injury litigation. *The Journal of Cognitive Rehabilitation*, *17*, 4–9.

Jing, E. T., Slick, D. J., Strauss, E., & Hultsch, D. F. (2002). How'd they do it? Malingering strategies on symptom validity tests. *Clinical Neuropsychologist*, *16*, 495.

Larrabee, G. J. (2000). Neuropsychology in personal injury litigation. *Journal of Clinical and Experimental Neuropsychology*, *22*, 702–707.

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.

Messinis, L., Tsakona, I., Malefaki, S., & Papathanasopoulos, P. (2007). Normative data and discriminant validity of Rey's Verbal Learning Test for the Greek adult population. *Archives of Clinical Neuropsychology*, *22*, 739–752.

Reitan, R. M. (1992). *Trail Making Test Manual for administration and scoring.* Tucson, AZ: Reitan Neuropsychological Laboratory.

Schmidt, M. (1996). *Rey Auditory Verbal Learning Test: A handbook*. Los Angeles, CA: Western Psychological Services.

Stevens, A., Friedel, E., Mehren, G., & Merten, T. (2008). Malingering and uncooperativeness in psychiatric and psychological assessment: Prevalence and effects in a German sample of claimants. *Psychiatry Research*, *157*, 191–200.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests* (3rd ed.). New York: Oxford University Press.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., & Jolles, J. (2005). Rey's verbal learning test: Normative data for 1855 healthy participants aged 24–81 years and the influence of age, sex, education, and mode of presentation. *Journal of the International Neuropsychological Society*, *11*, 290–302.

Wechsler, D. (1997). *WMS-III Wechsler Memory Scale* (3rd ed.). San Antonio: The Psychological Corporation.