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

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

WIM VAN DER ELST,[1] MARTIN P. J. VAN BOXTEL,[1] GERARD J. P. VAN BREUKELEN,[2]
AND JELLE JOLLES[1]

[1]Maastricht Brain and Behavior Institute, and European Graduate School of Neuroscience (EURON), Maastricht University
[2]Department of Methodology and Statistics, Maastricht University

Abstract

The Verbal Learning Test (VLT; Rey, 1958) evaluates the declarative memory. Despite its extensive use, it has been difficult to establish normative data because test administration has not been uniform. The purpose of the present study was to gather normative data for the VLT for a large number ($N = 1855$) of healthy participants aged 24–81 years, using a procedure in which the words to be learned were presented either verbally or visually. The results showed that VLT performance decreased in an age-dependent manner from an early age. The learning capacity of younger *versus* older adults differed quantitatively rather than qualitatively. Females and higher educated participants outperformed males and lower educated participants over the entire age range tested. Presentation mode affected VLT performance differently: auditory presentation resulted in a better recall on Trial 1 (a short-term or working memory measure), whereas visual presentation yielded a better performance on Trial 3, Trial 4, and Delta (a learning measure). (*JINS*, 2005, *11*, 290–302.)

INTRODUCTION

The declarative aspect of memory—the memory for specific facts or experiences (Squire, 1987)—is of particular interest in neuropsychology (Lezak, 1995; Woodruff-Pak, 1997). Consequently, a large number of tests to assess declarative memory have been developed, most notably the Verbal Learning Test of Rey (1958), the Buschke Selective Reminding Test (Buschke, 1973), and the California Verbal Learning Test (Delis et al., 1987). The present study focuses on the Verbal Learning Test (VLT; Brand & Jolles, 1985; Rey, 1958), an easily administered multitrial learning test (Lezak, 1995). In the VLT, fifteen monosyllabic words are presented in five subsequent trials, with a free recall procedure immediately following each presentation. After a delay of about 20 min, there is an additional free recall trial, followed by a recognition trial. Although the VLT is widely used in clinical practice and in memory research, its administration has not been uniform, which makes the finding of normative data difficult (Schmidt, 1996). Thus in this study, we sought to establish normative data for the VLT, using a procedure in which words were presented visually, on a computer screen, or auditorily, by means of an audio system. The normative data were derived from a cross-sectional study involving a large number of healthy participants ($N = 1855$) aged 24–81 years, matched for education and sex.

We decided to evaluate both auditory and visual stimulus presentation because earlier studies with other learning tasks suggested that presentation mode significantly affects the performance on short-term memory tasks. More specifically, auditorily presented material is recalled better than visually presented material. This phenomenon has been termed the "modality effect" (see e.g., Greene, 1992; Penney, 1989), and has been observed in both younger and

older adults (Taub, 1972, 1973). In contrast with short-term memory tasks, modality effects are conspicuously absent in long-term memory tasks (Penney, 1989). These results suggest that certain VLT measures, especially the Trial 1 score, might be significantly influenced by the mode of presentation, since it measures short-term memory. It is important to have normative data for two different sensory modality versions of the VLT (Van Boxtel et al., 2000) because participants, and especially older individuals, may have visual or hearing impairments. This is especially relevant in the clinical assessment of older patients (La Rue, 1992), in whom a diagnosis of dementia may depend—in part—on the VLT performance score.

Previous research has suggested that the VLT performance is significantly affected by demographic variables, most notably age, sex, and education. For a comprehensive overview of previous VLT research, the reader is referred to Schmidt (1996). Geffen et al. (1990) found in a study of 153 healthy participants (age range 16–86 years) that females outperformed males and that general VLT performance deteriorated with increasing age. Lannoo and Vingerhoets (1997) reported similar findings in their study of 200 healthy participants aged 18–74 years. They found that females, young adults, and participants with a higher educational level performed better than males, older adults, and lower educated participants. Uchiyama and colleagues (1995) investigated the influence of test form, age, ethnicity, and education on VLT performance in a large sample of 2056 healthy bisexual and homosexual HIV-seronegative males. Age was negatively and education positively related to general VLT performance. In contrast, Bolla-Wilson and Bleecker (1986) found that age affected performance on the first two learning trials only and that education did not have any effect in their study of 114 participants aged between 40–84 years. However, higher verbal intelligence was associated with a better performance as was sex, with females performing better than males. Savage and Gouvier (1992) investigated the influence of age and sex in a sample of 134 healthy participants aged 16–76 years. They reported a significant negative influence of increasing age on general VLT performance, whereas sex had no influence on test performance.

Since most—but not all—VLT studies suggested that demographic factors, such as age, sex, and education, significantly affect VLT performance, the influence of these variables on VLT performance was investigated (next to presentation mode) in the present study in order to adequately stratify the normative data.

METHOD

Participants

The data used in this study were derived from the Maastricht Aging Study (Bosma et al., 2003; Jolles et al., 1995; Van Boxtel et al., 1998, 2000), a prospective study of determinants of cognitive aging. Participants were recruited from the Registration Network Family Practices, a sample frame for research in primary care (Metsemakers et al., 1992). Individuals were invited to participate in the study by their general practitioners, rather than by the project staff, because this was expected to improve participation and study compliance. All participants were aged between 24–81 years, and all were without documented medical conditions that could interfere with normal cognitive aging: individuals with chronic neurological pathology (e.g., dementia, cerebrovascular disease, epilepsy, parkinsonism, and malignancies related to the nervous system), mental retardation, psychopathology, or chronic psychotropic drug use were excluded. The sample was stratified for age (12 levels, ranging from 25 ± 1 year, 30 ± 1 year, . . . , to 80±1 year), sex, and level of education (3 levels: low, average, and high). Level of education (LE) was assessed by classifying formal schooling in a system often used in the Netherlands (De Bie, 1987), which is comparable with the International Standard Classification of Education (UNESCO, Paris, 1976). Three groups were formed—those with at most primary education (LE low), those with junior vocational training (LE average), and those with senior vocational or academic training (LE high). These three levels of education corresponded with an average of respectively 8.60, 11.41, and 15.25 years of full-time education in the sample ($SD =$ 1.95, 2.48, and 3.31, respectively).

Not all data for the 1855 participants administered the VLT were included in the analyses. Data of participants who scored less than 24 on the Mini-Mental State Examination (Folstein et al., 1975) were excluded, since they could not be considered to be "cognitively intact". For the auditory VLT version, an additional exclusion criterion was a hearing loss of 35 dB or more in the better ear. After these corrections—which excluded the data of 75 participants— the data of 1780 participants were analyzed (auditory VLT, $n = 450$; visual VLT $n = 1330$). Basic descriptive data for the sample are provided in Table 1. The ethnic background of all participants was Caucasian, and all participants were native Dutch speakers.

Procedure and Instruments

All participants were tested individually at the neuropsychological laboratory of the Brain and Behaviour Institute (the Netherlands) using a Dutch adaptation of the VLT, the *Groningen Vijftienwoorden Test* (Groningen Fifteen Words Test), which had been developed earlier by Brand and Jolles (1985). These authors constructed six VLT versions using a pool of 410 monosyllabic words in the Dutch language. These words have a frequency of 20 to 400 per million according to the De La Court count of Dutch words (Linschoten, 1963). Only common words were used which are acquired early in life (before age 6 years) and refer to concrete objects such as flower, crown, bear, tree, etc. Potentially ambiguous words and words referring to emotional factors were not used. Various possible categories of words

Table 1. Descriptive characteristics of the sample ($N = 1780$)

Age group	N	Age Mean	Age SD	Level of education frequency Low	Level of education frequency Average	Level of education frequency High	Male:female ratio
25 ± 1 year	157	25.43	.90	21	75	61	78:79
30 ± 1 year	151	30.62	.87	20	78	53	77:74
35 ± 1 year	159	35.54	.91	34	76	49	79:80
40 ± 1 year	156	40.51	.88	36	74	46	77:79
45 ± 1 year	161	45.54	.92	50	71	40	80:81
50 ± 1 year	161	50.30	.90	53	71	37	79:82
55 ± 1 year	161	55.44	.90	84	50	27	82:79
60 ± 1 year	158	60.60	.83	73	65	19	82:76
65 ± 1 year	158	65.47	.87	86	52	20	78:80
70 ± 1 year	153	70.29	.82	82	51	20	81:72
75 ± 1 year	147	75.14	.78	78	47	21	69:78
80 ± 1 year	58	79.83	.87	36	8	13	25:33
Total	1,780	51.29	16.22	653	718	406	887:893

Note. Data on level of education were missing from 3 participants.

(e.g., animals, body parts, parts of the house, furniture) were evenly distributed over the lists in order to control for potential semantic associations. Auditory associations were controlled for each list. The test versions were evaluated in control experiments after which particular words were discarded and replaced by new words. The test versions were eventually shown to be parallel and used in numerous drug studies in which repeated administration of parallel versions of tests is essential (e.g., Hijman et al., 1992). In the present study, we used one of these six parallel versions.

The test instructions were read aloud by the experimenter. The participant was required to listen to, or to watch, the fifteen words that were presented by means of an audio system (auditory version) or on a computer screen (visual version) and to repeat as many words as possible as soon as the presentation stopped. There was no restriction on the output order. The first trial was followed by four more trials in which the words were presented in identical order. When the fifth trial was completed, a fixed battery of other cognitive tests was administered for about 20 min. These tests did not involve the learning of verbal material in order to avoid interference with the previously learned VLT words [for a description of these tests, see Jolles et al. (1995)]. After the delay—and unexpectedly for the participants— the instruction was given to recall the words learned (delayed recall). This was followed immediately by a recognition test, involving yes/no recognition of the fifteen words intermixed with fifteen nontarget words.

Stimulus words were presented for 1 s, with an interstimulus interval of 1 s. In the auditory VLT, the words were presented by the digitized voice of a professional speaker, using a high-fidelity stereo speaker system connected to a standard PC-microcomputer. In the visual VLT, stimuli were presented on a 17-inch computer screen operated by a standard PC-microcomputer. All words were displayed in capital letters of the font 'Times' size 72, in white against a black background and in the center of the screen. If necessary, the visual VLT was performed with optical correction for nearby vision.

The auditory VLT version was administered to one-fourth of the sample, and the visual VLT to three-fourths of the sample. The participants were divided at random over the conditions on a 1:3 basis (ratio of the number of participants in the auditory *versus* visual VLT condition) and with the restriction that the proportion of participants that were assessed with the visual *versus* auditory VLT was equal for each Age group × Level of education × Sex cell.

The data were collected by five test assistants who had been intensively trained in test administration by the neuropsychologists and physician of the project staff. Members of the project staff visited the test assistants during assessment of the VLT at least once a week to ensure uniform test administration and data collection.

Outcome Measures

The following measures were analyzed: (1) Trial 1, 2, 3, 4, and 5 score, the number of correctly recalled words in each of these learning trials. (2) Repetitions, that is, the total number of words repeated over Trials 1–5. Only words that were repeated during the first recall-attempt per trial were scored as repetitions (e.g., flower–crown–bear–crown–. . .). This precludes that words repeated during the often-seen strategy of overtly repeating a sequence of words are scored as repetitions. (3) Delayed recall, the number of correctly recalled words after the 20-min delay. (4) Total recall 1–5, the total number of correctly recalled words summed over the five learning trials. (5) Total recall 1–3, the total number of correctly recalled words summed over the first three learning trials. (6) Delta, which was calculated as the Trial 3 score minus the Trial 1 score. The latter two scores were used because ceiling effects are often

observed in the VLT, especially in younger adults: the increase in the number of words recalled is especially pronounced during the first three learning trials, which makes the scores for the last two learning trials to some extent redundant. Delta differs from the Total recall 1–3 score because it is a learning measure that corrects for the Trial 1 score, which is a measure of short-term and working memory. (7) Recognition, calculated as the number of correct hits minus the number of false positives. This measure is equivalent to the Recognition Percent Correct score (= [true positives + true negatives]/30) proposed by Harris et al. (2002). (8) A Retention score, calculated as the Delayed recall score divided by the Trial 5 score. Thus a retention score of .75 indicates that the individual's score on the delayed recall score was 25% lower than that person's score in Trial 5.

Data Analysis

Multiple linear regression models were fitted using a stepdown hierarchical procedure for the VLT measures adjusted for Age, (Age)^2 (which enables modelling of quadratic age effects), Sex, Level of education, and Mode of presentation. Since previous VLT research provided equivocal results with respect to the influence of demographic factors on VLT performance—especially with respect to education and sex—the interactions between Age and Education, Age and Sex, (Age)^2 and Education, and (Age)^2 and Sex were also included in the model. Age was centered (Age = calendar age minus 50) before the quadratic terms and interactions were computed. This was done in order to avoid multicollinearity (Marquardt, 1980; Neter et al., 1989), because high levels of multicollinearity may lead to technical problems when estimating the regression coefficients (Aiken & West, 1991). Level of education (LE) was dummy coded with LE average as the reference category. The dummies LE low and LE high were always either both included in, or both excluded from, the model, since they belong together and represent the effect of the categorical predictor education. Similarly, their interactions with another predictor were always either both included in or excluded from the model. Nonsignificant predictors ($p > .005^*$) were excluded from the model, but no predictor was removed from the model as long as it was also included in a higher order term in the model. In particular, Age was never removed if (Age)^2 or any interaction involving Age or (Age)^2 was still in the model. The reason for this is that the p-value of any predictor is arbitrary (depending on the coding used for predictors) if that predictor is part of a higher order predictor in the model (Aiken & West, 1991). The assumptions of regression analysis (homoscedasticity, normal distribution of the residuals, absence of multicollinearity, and absence of "influential cases") were tested for each model. Homoscedasticity was evaluated by visual inspection of the scatter plots of the

---

*A lower *alpha* level was chosen in order to avoid Type I errors due to multiple testing.

residuals against the predicted values. Normal distribution of the residuals was investigated by visual inspection of the histograms and the normal probability plots. The occurrence of multicollinearity was checked by calculating the Variance Inflation Factors (VIFs), which should not exceed 10 (Belsley et al., 1980). Cook's distances were calculated to identify possible influential cases. All analyses were performed using the SPSS 10.0 for Macintosh software package.

Normative data can then be obtained by calculating the residuals for each VLT score (e_i = observed score − predicted score): after standardization of the residuals ($Z_i = e_i/SD$[residual]), the performance of the participant can be evaluated *via* a Z distribution table with cumulative probabilities.

RESULTS

The regression models are presented in Table 2. No significant influence of outliers was observed (maximum Cook's distance .07). The VIFs of the predictors in the regression models were at most 2.7, well below the cut-off value. The Repetitions score was square root transformed ($\sqrt{\text{Repetitions}}$) because of positively skewed residuals before transformation of the score.

Age was by far the strongest predictor of performance on all VLT scores (see standardized B's, Table 2), with the exception of Delta. An additional quadratic effect of age on VLT performance was observed for most VLT scores, with the exception of Delta, although the quadratic trend was less important than the linear age effect. Females and highly educated participants scored significantly higher than males and their less educated counterparts on most VLT measures. Mode of presentation differentially affected VLT performance: the Trial 1 score was higher with the auditory VLT compared to the visual VLT, and the opposite was found for the Trial 3, 4, Delta, and $\sqrt{\text{Repetitions}}$ score. The influence of presentation mode on the VLT Trials 1–5 and the Delayed recall trial performance is shown in Figure 1, for four broad age groups (24–36 years, 39–51 years, 54–66 years, and 69–81 years).

Only one significant interaction was found: an Age × LE low interaction on the Trial 1 score. Figure 2 depicts the predicted Trial 1 scores for male participants for the visual VLT (the shape of these plots was identical for females and for the auditory VLT, with a constant value added or subtracted to the predicted scores—see Table 2). As shown in Figure 2, low-educated participants performed worse compared to average/high-educated participants for all ages, but the difference in performance between the level of education groups decreased with increasing age.

Normative data are obtained by calculating the participant's predicted VLT scores, using the regression models (see Table 2). The residuals of each score are then calculated (e_i = observed score − predicted score), and standardized ($Z_i = e_i/SD$[residual]) using Table 3. Note that the scatter plots of the residuals against the predicted values showed a trend to heteroscedasticity (a decrease in

Table 2. Multiple linear regression models of the VLT scores following a step-down hierarchical procedure: the full model included Age, (Age)^2, LE low, LE high, Sex, Mode of Presentation, Age × LE low, Age × LE high, Age × Sex, (Age)^2 × LE low, (Age)^2 × LE high and (Age)^2 × Sex

Score	Variable	B	Std. Error B	T	Standardized B	R^2
Trial 1	(constant)	5.627	.085	66.323**		
	Age	−.048	.004	−12.644**	−.422	
	(Age)^2	−.0005	0	−2.959*	−.064	
	Sex	−.425	.076	−5.578**	−.116	
	LE low	−.537	.094	−5.745**	−.141	
	LE high	.417	.102	4.092**	.095	
	Mode of presentation	.353	.087	4.059**	.083	
	Age × LE low	.018	.006	3.057*	.095	
	Age × LE high	−.0074	.006	−1.187	−.032	.259
Trial 2	(constant)	8.972	.098	91.097**		
	Age	−.052	.003	−17.113**	−.378	
	(Age)^2	−.0006	0	−2.899*	−.060	
	Sex	−.808	.092	−8.751**	−.181	
	LE low	−.594	.108	−5.483**	−.128	
	LE high	.514	.121	4.261**	.097	.257
Trial 3	(constant)	10.909	.109	100.340**		
	Age	−.053	.003	−16.510**	−.363	
	(Age)^2	−.0008	0	−3.771**	−.078	
	Sex	−.971	.098	−9.904**	−.204	
	LE low	−.678	.115	−5.890**	−.137	
	LE high	.491	.128	3.833**	.087	
	Mode of presentation	−.493	.112	−4.408**	−.090	.265
Trial 4	(constant)	11.865	.107	111.229**		
	Age	−.050	.003	−15.915**	−.349	
	(Age)^2	−.0009	0	−4.776**	−.099	
	Sex	−.995	.096	−10.343**	−.213	
	LE low	−.691	.113	−6.121**	−.142	
	LE high	.523	.126	4.157**	.094	
	Mode of presentation	−.358	.110	−3.266**	−.066	.265
Trial 5	(constant)	12.462	.103	121.489**		
	Age	−.046	.003	−14.632**	−.327	
	(Age)^2	−.0007	0	−3.640**	−.077	
	Sex	−1.018	.096	−10.584**	−.222	
	LE low	−.670	.113	−5.932**	−.141	
	LE high	.497	.126	3.956**	.091	.237
Total recall 1–5	(constant)	49.672	.401	123.945**		
	Age	−.247	.012	−19.916**	−.416	
	(Age)^2	−.0033	.001	−4.250**	−.084	
	Sex	−4.227	.376	−11.253**	−.220	
	LE low	−3.055	.441	−6.925**	−.153	
	LE high	2.496	.491	5.081**	.109	.334
Total recall 1–3	(constant)	25.440	.243	104.816**		
	Age	−.150	.007	−19.968**	−.422	
	(Age)^2	−.0016	0	−3.454**	−.069	
	Sex	−2.217	.227	−9.748**	−.193	
	LE low	−1.699	.267	−6.361**	−.143	
	LE high	1.467	.297	4.932**	.107	.318
Delta	(constant)	5.100	.068	74.881**		
	Age	−.013	.003	−4.747**	−.109	
	Sex	−.502	.089	−5.651**	−.130	
	Mode of presentation	−.858	.102	−8.385**	−.192	.064

Table 2. Continued

Score	Variable	B	Std. Error B	T	Standardized B	R^2
Delayed recall	(constant)	10.924	.128	85.374**		
	Age	−.073	.004	−18.440**	−.401	
	$(Age)^2$	−.0009	0	−3.704**	−.076	
	Sex	−1.197	.120	−9.985**	−.203	
	LE low	−.844	.141	−5.991**	−.138	
	LE high	.424	.157	2.707	.060	.282
Retention	(constant)	.871	.007	125.526**	`	
	Age	−.003	0	−13.652**	−.310	
	$(Age)^2$	−.00005	0	−3.405*	−.077	
	Sex	−.029	.008	−3.812**	−.085	.116
Recognition	(constant)	14.191	.090	157.143**		
	Age	−.034	.003	−10.757**	−.248	
	$(Age)^2$	−.0008	0	−3.867**	−.089	
	Sex	−.398	.101	−3.957**	−.090	.084
√Repetitions	(constant)	2.052	.040	50.947**		
	Age	−.013	.001	−9.815**	−.226	
	$(Age)^2$	−.0004	0	−3.945**	−.091	
	Sex	−.129	.043	−3.021*	−.068	
	Mode of presentation	−.280	.049	−5.702**	−.129	.086

Note. LE = Level of education. Coding of the predictors: Age = calendar age − 50; $(Age)^2$ = (calendar age − 50)^2; Sex: Male = 1, Female = 0; LE Low: Low Education = 1, Average or High Education = 0; LE High: High Education = 1, Low or Average Education= 0; Mode of Presentation: Auditory version = 1, Visual version = 0
*$p < .005$; **$p < .001$.

the variance as a function of the predicted scores) for the Recognition and Retention scores. Therefore, the standard deviations of the residuals of these two scores were calculated as a function of the predicted scores divided into quartiles (Table 3). After standardization of the residuals, the participant's performance can be evaluated *via* a Z distribution table with cumulative probabilities (or a simplified version thereof, see Table 4).
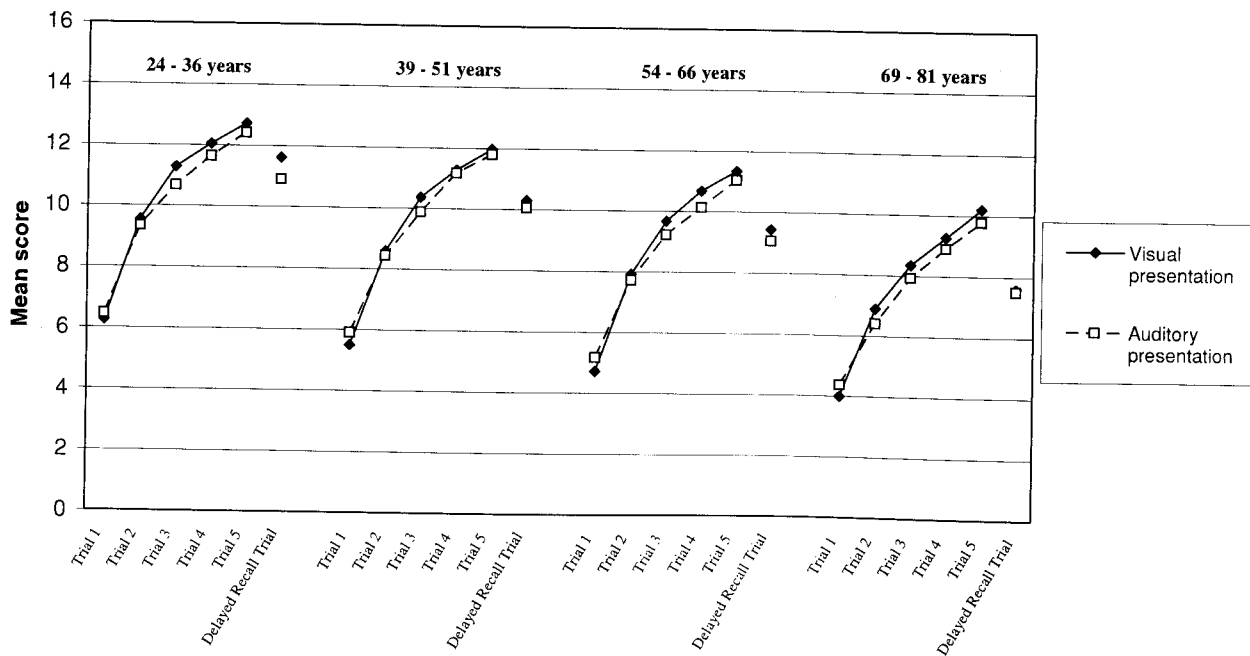

Fig. 1. Mean VLT Trial 1, 2, 3, 4, 5, and Delayed recall scores for four Age groups (24–36 years, 39–51 years, 54–66 years, and 69–81 years), separated for Presentation mode.
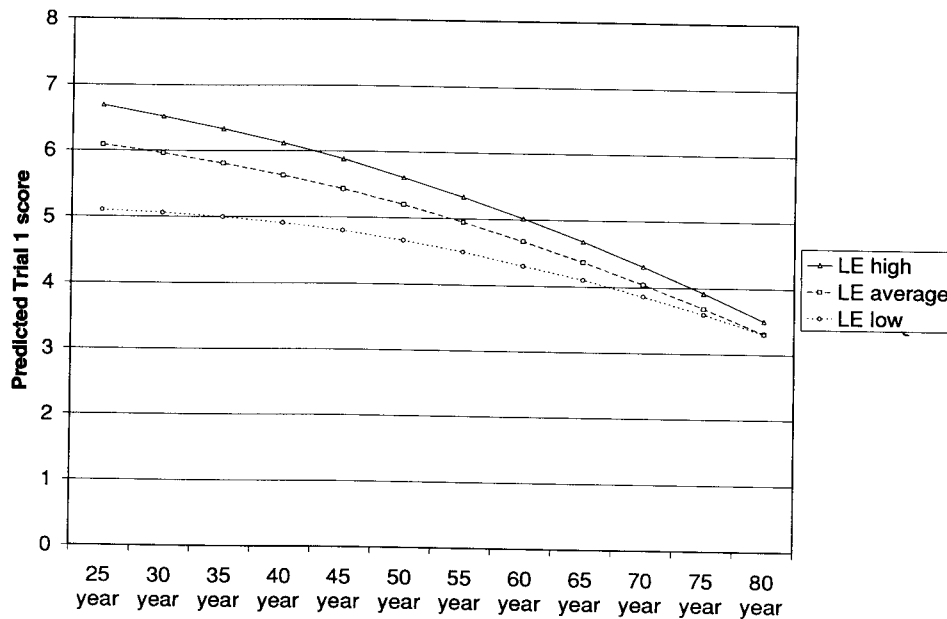
Fig. 2. Predicted VLT Trial 1 scores for male participants on the visual VLT, per Level of education category as a function of Age.

For example, the visual VLT was administered to a 60-year-old man with a low level of education. The participant recalled 5 words in the Delayed recall trial. The predicted Delayed recall score for this person would be 8.063 $(= 10.924 + ((60 - 50) * -.073) + ((60 - 50)^2 * -.0009) + (1 * -1.197) + (1 * -.844) + (0 * .424))$, with a residual of -3.063 $(e_i = 5 - 8.063)$. The SD(residual) equals 2.496 and the standardized residual -1.227 $(= -3.063/2.496)$,

Table 3. Standard deviations of the residuals per VLT measure

Measure	Predicted score	SD (residual)
Trial 1	General (all scores)	1.584
Trial 2	General (all scores)	1.923
Trial 3	General (all scores)	2.042
Trial 4	General (all scores)	2.003
Trial 5	General (all scores)	2.003
Total recall 1–5	General (all scores)	7.826
Total recall 1–3	General (all scores)	4.739
Delta	General (all scores)	1.873
Delayed recall	General (all scores)	2.496
Retention	General (all scores)	.163
	≤.795	.191
	Between .796 and .852	.172
	Between .853 and .886	.154
	≥.887	.127
Recognition	General (all scores)	2.119
	≤13.285	2.972
	Between 13.286 and 13.937	2.250
	Between 13.938 and 14.172	1.549
	≥14.172	1.300
√Repetitions	General (all scores)	.899

which corresponds to a p-value of .11, and can be considered "below normal" according to Table 4.

Tables A1–A4 (in Appendix) provide normative tables based on the regression models in Table 2 for the Trial 1, Delta, Delayed recall, and Recognition scores, stratified by their statistically significant predictors.† If an individual is not exactly 25, 30, . . . , 80 years old, then the person's age should be rounded up to the closest age given. If the performance of such an individual is borderline-normal according to these tables (e.g., Z value ≈ -1.28), the regression models presented in Table 2 can be used to determine the exact Z values.

DISCUSSION

The results of the present study showed highly significant effects of age on VLT performance. Interestingly, the age-related decline in performance started early in life, in a highly linear fashion: the linear age effect was a much stronger predictor of VLT performance than the quadratic age effect. Put in another way, the age-related differences in general VLT performance were more manifest in a quantitative than in a qualitative way—although the absolute number of recalled words differed as a function of age, the shape of the learning curve was similar (see Figure 1). The results suggested that younger adults outperformed older adults because of a better performance on Trial 1. These

---

†Due to space constraints, we only provide normative tables for the most important VLT measures here. Other scores can be evaluated using the described normative procedure. In the course of 2005, a normative handbook will also made available *via* Neuropsych Publishers (http://www-np.unimaas.nl/)

Table 4. Z score intervals with the corresponding proportion of individuals with a score falling within this interval and corresponding verbal labels

Z Score interval	Percentage within this interval	Verbal label
Below -2.0	2.3%	Very poor performance
Between -2.0 and -1.5	4.4%	Poor performance
Between -1.5 and -1.0	9.2%	Below normal
Between -1.0 and -0.5	15%	Normal
Between -0.5 and 0.5	38.3%	Normal
Between 0.5 and 1.0	15%	Normal
Between 1.0 and 1.5	9.2%	Above normal
Between 1.5 and 2.0	4.4%	Good performance
Above 2.0	2.3%	Very good performance

differences in Trial 1 performance may be due to differences in processing speed, as suggested by Bolla-Wilson and Bleecker (1986). In clinical practice, the Total recall 1–5 score is the most commonly used learning summary score (Schmidt, 1996). The results of the present study, however, suggest that this measure is not the most appropriate one. Such a summary measure suggests that the differences in learning capacity over the first five trials in the VLT vary largely as a function of age, which was not supported by the results of the present study (see Figure 1). The same criticism applies to the Total recall 1–3 score, although this score has the advantage over the Total recall 1–5 score in that it takes ceiling effects, at least theoretically, into account. A measure such as Delta (= Trial 3 − Trial 1 score) might be a more appropriate learning measure, since it corrects for Trial 1 performance and takes ceiling effects into account, and thus gives a more realistic estimation of the verbal learning capacity of an individual.

Although Schmidt (1996) suggested that the effect of sex appears to be of limited importance in interpreting VLT results, the present study clearly showed significant effects of sex on all VLT measures, with women consistently outperforming men. This finding is in agreement with studies that point to the superior verbal abilities of women, for example on verbal episodic memory tasks (e.g., Herlitz & Yonker, 2002; Yonker et al., 2003) and verbal fluency tasks (e.g., Rideout & Winchester, 1990). With regard to education, the negative impact of a low level of education—as compared to an average level of education—on VLT performance was larger than the positive impact of a high level of education on most VLT scores (see the standardized B values in Table 2). Furthermore, the significant Age × Low level of education interaction on the Trial 1 score suggested that the differences in performance between the low level of education group and the average/high level of education group decreased with increasing age.

It should be noted that Bolla-Wilson and Bleecker (1986) found no effect of education on VLT performance, in contrast to the results of the present study. This difference is most likely attributable to the fact that Bolla-Wilson and Bleecker (1986) used education together with a verbal intelligence measure as predictors of the VLT performance (next to age and sex). It is possible that a variable such as (verbal) intelligence is a more appropriate estimator of a person's general ability than is education, especially in older individuals. Indeed, before the 1960s most people did not continue their schooling, mostly for reasons other than their intellectual capacity, for example, for socio-economic reasons (Jolles et al., 1995). Nevertheless, we used level of education as an estimator of general ability instead of verbal intelligence for two main reasons. First, in order to facilitate the use of normative data, it is important to use a readily available stratification variable. Use of verbal intelligence as a predictor of the VLT performance means that a person would also have to complete a verbal intelligence test before the normative VLT data could be used. This is time-consuming and expensive, and might be difficult in certain groups, for example, elderly people, people with attention disorders, people with a hearing impairment, etc. Second, as with any measurement of a psychological function, there is an error component in the obtained score of a verbal intelligence test. This would introduce an extra source of error to the model (the error component in the score on the verbal intelligence test) in addition to the error component for the different VLT measures.

Mode of stimulus presentation had a significant, but not uniform, influence on some VLT measures. Thus the Trial 1 score was higher with the auditory VLT (in accordance with the modality effect), the Trial 3, 4, Delta, and $\sqrt{}$ Total repetition scores were higher with the visual VLT. It should be noted that in the models predicting the Delayed recall and Trial 5 scores, the predictor Mode of presentation was deleted in the final step of the hierarchical procedure, but was almost significant ($p = .008$ and $p = .010$, respectively). Thus, the pattern of results suggests that auditory presentation is associated with a better performance of tasks depending mainly on short-term or working memory processes (VLT Trial 1), whereas visual presentation is associated with a better performance of tasks that depend on long-term memory processes (VLT Delayed recall trial) or tasks that depend on a mixture of short-term and long-term memory processes (VLT Trials 3, 4, and 5). This stresses the importance of using normative data appropriate for the test procedure (stimulus presentation) used. This aspect tends to be overlooked in clinical practice, when tests are computerized even though the normative data were obtained for noncomputerized tests.

At present, it is not known whether the visual VLT or the auditory VLT version should be used as first choice to assess verbal memory or to detect pathological conditions such as, for example, dementia. This issue merits further investigation. However, an important issue to keep in mind regarding this question concerns earlier findings by Van Boxtel and colleagues (2000). These authors suggested that the auditory VLT version underestimates the true memory capacity of people with mild to moderate degrees of hearing loss, even after the volume at which the VLT words are pre-
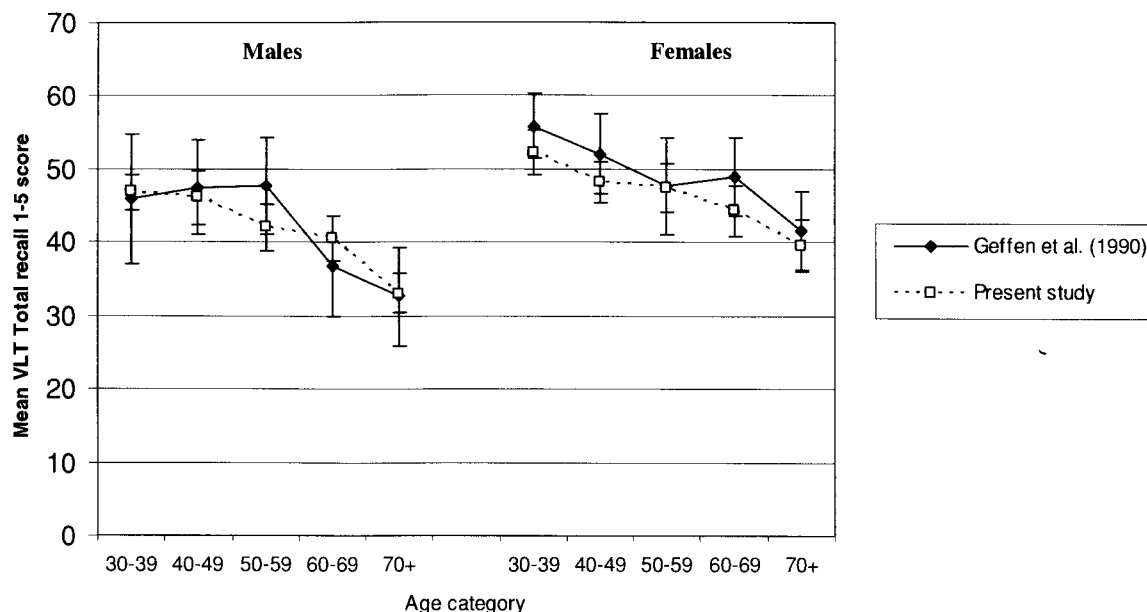
Fig. 3. Mean auditory VLT Total recall 1–5 scores for males (left) and females (right) separated for Age category (and 95% Confidence Interval).

sented is adjusted for individual hearing ability. Importantly, Van Boxtel et al. (2000) also reported a large overlap in hearing acuity between people who did or did not report hearing loss. Therefore, we recommend that the visual VLT version be used to test participants with an unknown (objective) hearing status (even if the participant claims not to have a hearing loss) or with a (even mild) hearing impairment.

The present data were obtained for a healthy population of native Dutch speakers. It can be questioned whether the normative data of the Dutch VLT are applicable to individuals who have a different native language or cultural background. Recent studies suggested that words that are selected using the same criteria (e.g., comparable with regard to semantic and phonemic features, words that seem equally concrete, words that seem as difficult or easy to cluster by association, etc.) lead to comparable test versions across countries and languages (Van den Burg & Kingma. 1999). This was the case when we compared our data for the auditory VLT version with the results of Geffen and colleagues (1990) obtained with an English-speaking sample. After redefining the age categories of our sample to match those of Geffen et al. (1990) (see Figure 3), we found that the mean VLT Trial 1, 2, 3, 4, 5, and Total recall 1–5 scores in our sample were not significantly different from the corresponding mean scores in their sample (all $p > .01$‡) for all age by sex groups (see Figure 3 for the Total recall 1–5 score).

We were not able to compare the scores for the visual VLT (because Geffen and colleagues did not administer a

visual VLT version), and for certain VLT measures, such as Recognition and Delayed recall (because Geffen and colleagues used different test administration and scoring procedures, which means that these scores are not comparable). However, we expect it to be unlikely that English-speaking people would score differently from Dutch-speaking people on a visual version of the VLT, especially when differences were not found with the auditory version of the VLT. Supporting evidence for this claim can also be found in a study of Møller and colleagues (1998), who showed that the Visual Verbal Learning Test (in three trials instead of five) was culturally robust when evaluated in seven different language areas (English, Danish, Dutch. French, German, Greek, and Spanish). Thus, the results of these studies suggest that normative data for a Dutch sample (and normative data for a sample with a language very similar to Dutch, such as German; see Heubrock, 1994) can also be used for samples of English-speaking people.

In summary. we established normative data for the auditory and visual version of the VLT. The data were derived from a large sample of 1855 cognitively intact men and women of different educational levels and ages (range 24–81 years). We used multiple regression models to establish normative data that account for the effects of these variables. The results showed that VLT performance is superior in women, decreases with age, and is associated with level of education. Performance on VLT Trial 1, which depends mainly on short-term or working memory, was better with the auditory version, whereas performance on VLT trials that depend on long-term memory processes or trials that depend on a combination of short-term and long-term memory processes was better with the visual version. Furthermore, although the Total recall 1–5 score is currently the most frequently used VLT learning summary measure, our

---

‡A series of two-tailed *t*-tests per age by sex group were performed instead of a more appropriate technique, for example ANOVA with Age category and Sex as between subject factors, since we did not have access to the raw dataset of Geffen et al. (1990).

results indicated that Delta (= Trial 3 − Trial 1) may provide a better estimation of the verbal learning capacity of an individual.

REFERENCES

Aiken, L.S. & West, S.G. (1991). *Multiple regression: Testing and interpreting interactions.* Newbury Park, California: Sage.

Belsley, D.A., Kuh, E., & Welsch, R.E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity.* New York: Wiley.

Bolla-Wilson, K. & Bleecker, M. L. (1986). Influence of verbal intelligence, sex, age, and education on the Rey auditory verbal learning test. *Developmental Neuropsychology, 2,* 203–211.

Bosma, H., Van Boxtel, M.P.J., Ponds, R.W.H.M., Houx, P.J., & Jolles, J. (2003). Mental work demands protect against cognitive impairment: MAAS prospective cohort study. *Experimental Aging Research, 29,* 33–45.

Brand, N. & Jolles, J. (1985). Learning and retrieval rate of words presented auditory and visually. *Journal of General Psychology, 112,* 201–210.

Buschke, H. (1973). Selective reminding analysis of memory and learning. *Journal of Verbal Learning and Verbal Behaviour, 12,* 543–550.

De Bie, S.E. (1987). *Standaardvragen 1987: Voorstellen voor uniformering van vraagstellingen naar achtergrondkenmerken en interviews* [Standard questions 1987: Proposal for uniformization of questions regarding background variables and interviews]. Leiden, The Netherlands: Leiden University Press.

Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (1987). *California Verbal Learning Test Manual.* San Antonia, Texas: The Psychological Corporation.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Geffen, G., Moar, K.J., O'Hanlon, A.P., Clarck, C.R., & Geffen, L.N. (1990). Performance measures of 16- to 86-year-old males and females on the Auditory Verbal Learning Test. *The Clinical Neuropsychologist, 4,* 45–63.

Greene, R.G. (1992). *Human memory.* Hove, UK: Lawrence Erlbaum Associates Ltd.

Harris, M.E., Ivnik, R.J., & Smith, G.E. (2002). Mayo's Older Americans normative studies: Expanded AVLT recognition trial norms for ages 57 to 98. *Journal of Clinical and Experimental Neuropsychology, 24,* 214–220.

Herlitz, A. & Yonker, J.E. (2002). Sex differences in episodic memory: The influence of intelligence. *Journal of Clinical and Experimental Neuropsychology, 24,* 107–114.

Heubrock, D. (1994). Auditiv-verbales Lernen unter standardisierten Bedingungen: Erste deutsche Normen für 16- bis 26-jährige Männer und Frauen zum Auditiv-Verbalen Lerntest (AVLT) [Standardized auditory-verbal learning: Preliminary German Auditory-Verbal Learning Test (AVLT) norms for a male and female population aged 18 to 26]. *Zeitschrift für Differentielle und Diagnostische Psychologie, 15,* 65–76.

Hijman, R., Jolles, J., Verhoeven, W.M.A., Van Ree, J.M., Elderson, A., & De Wied, D. (1992). Desglycinamide-(Arg8)-vasopressin in five trials with memory-disturbed patients. *Human Psychopharmacology: Clinical and Experimental, 7,* 7–23.

Jolles, J., Houx, P.J., Van Boxtel, M.P.J., & Ponds, R.W.H.M. (1995). *Maastricht aging study: Determinants of cognitive aging.* Maastricht, The Netherlands: Neuropsych Publishers.

Lannoo, E. & Vingerhoets, G. (1997). Flemish normative data on common neuropsychological tests: Influence of age, education, and sex. *Psychologica Belgica, 37,* 141–155.

La Rue, A. (1992). *Aging and neuropsychological assessment.* New York: Plenum Press.

Lezak, M.D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Linschoten, J. (1963). *De la Court frequentie tellingen van Nederlandse woorden* (Report No. 6301). Rijks Universiteit Utrecht, Psychologisch Laboratorium.

Marquardt, D.W. (1980). You should standardize the predictor variables in your regression models. *Journal of the American Statistical Association, 75,* 87–91.

Metsemakers, J.F.M., Höppener, P., Knottnerus, J.A., Kocken, R.J.J., & Limonard, C.B.G. (1992). Computerized health information in the Netherlands: A registration network of family practices. *British Journal of General Practice, 42,* 102–106.

Møller, J.T., Cluitmans, P., Rasmussen, L.S., Houx, P.J., Rasmussen, H., Canet, J., Rabbitt, P., Jolles, J., Larsen, K., Hanning, C.D., Langeron, O., Johnson, T., Lauven, P.M., Kristensen, P.A., Biedler, A., van Beem, H., Fraidakis, O., Silverstein, J.H., Beneken, J.E.W., & Gravenstein, J.S. (1998). Long-term postoperative cognitive dysfunction in the elderly: ISPOCD1 study. *The Lancet, 351,* 857–861.

Neter, J., Wasserman, W., & Kutner, M.H. (1989). *Applied linear regression models* (2nd ed.). Homewood, Illinois: Irwin.

Penney, C.G. (1989). Modality effects and the structure of short-term verbal memory. *Memory and Cognition, 17,* 398–422.

Rey, A. (1958). *L'examin clinique en psychologie.* Paris, France: Presses Universitaires de France.

Rideout, B.E. & Winchester, L. (1990). Effects on verbal fluency associated with sex, practice, and stimulus letter. *Perceptual and Motor Skills, 71,* 656–658.

Savage, R.M. & Gouvier, W.D. (1992). Rey auditory-verbal learning test: The effects of age and gender, and norms for delayed recall and story recognition trials. *Archives of Clinical Neuropsychology, 7,* 407–414.

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

Squire, L.R. (1987). *Memory and brain.* London, UK: Oxford University Press.

Taub, H.A. (1972). A comparison of young adult and old groups on various digit span tasks. *Developmental Neuropsychology, 6,* 60–65.

Taub, H.A. (1973). Memory span, practice, and aging. *Journal of Gerontology, 28,* 335–338.

Uchiyama, C.L., D'Elia, L.F., Dellinger, A.M., Becker, J.T., Selnes, O.A., Welsch, J.E., Chen, B.B., Satz, P., Van Gorp, W., & Miller, E.N. (1995). Alternate forms of the auditory-verbal learning test: Issues of test comparability, longitudinal reliability, and moderating demographic variables. *Archives of Clinical Neuropsychology, 10,* 133–145.

United Nations Educational, Scientific and Cultural Organisation (UNESCO) (1976). *International Standard Classification of Education (ISCED).* Paris, France: UNESCO.

Van Boxtel, M.P.J., Buntinx, F., Houx, P.J., Metsemakers, J.F.M., Knottnerus, A., & Jolles, J. (1998). The relation between morbidity and cognitive performance in a normal aging popula-

tion. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 53A,* M147–M154.

Van Boxtel, M.P.J., van Beijsterveldt, C.E.M., Houx, P.J., Anteunis, L.J.C., Metsemakers, J.F.M., & Jolles, J. (2000). Mild hearing impairment can reduce verbal memory performance in a healthy adult population. *Journal of Clinical and Experimental Neuropsychology, 22,* 147–154.

Van den Burg, W. & Kingma, A. (1999). Performance of 225 Dutch school children on Rey's Auditory Verbal Learning Test

(AVLT): Parallel test–retest reliabilities with an interval of 3 months and normative data. *Archives of Clinical Neuropsychology, 14,* 545–559.

Woodruff-Pak, D.S. (1997). *The neuropsychology of aging.* Malden: Blackwell Publishers.

Yonker, J.E., Eriksson, E., Nilsson, L., & Herlitz A. (2003). Sex differences in episodic memory: Minimal influence of estradiol. *Brain and Cognition, 52,* 231–238.

APPENDIX

Table A1. Normative VLT data for the Trial 1 score stratified by Mode of presentation, Level of education, Sex, and, Age (25, 30, . . . , 80 years). The raw test score leading to a particular Z-value is given for Z-values indicating the percentiles 5, 10, 20, 50, 80, 90, and 95

		Z value	Cum. prob.	Male Age in years											Female Age in years											
				25	30	35	40	45	50	55	60	65	70	75	80	25	30	35	40	45	50	55	60	65	70	75	80
Visual VLT	LE low	1.64	.95	7.7	7.7	7.6	7.5	7.4	7.3	7.1	6.9	6.7	6.5	6.2	5.9	8.1	8.1	8.0	7.9	7.8	7.7	7.5	7.3	7.1	6.9	6.6	6.3
		1.28	.90	7.1	7.1	7.0	6.9	6.8	6.7	6.5	6.3	6.1	5.9	5.6	5.3	7.6	7.5	7.5	7.4	7.3	7.1	7.0	6.8	6.6	6.3	6.1	5.8
		.84	.80	6.4	6.4	6.3	6.2	6.1	6.0	5.8	5.6	5.4	5.2	4.9	4.6	6.9	6.8	6.8	6.7	6.6	6.4	6.3	6.1	5.9	5.6	5.4	5.1
		0	.50	5.1	5.1	5.0	4.9	4.8	4.7	4.5	4.3	4.1	3.9	3.6	3.3	5.5	5.5	5.4	5.3	5.2	5.1	4.9	4.7	4.5	4.3	4.0	3.7
		−.84	.20	3.8	3.7	3.7	3.6	3.5	3.3	3.2	3.0	2.8	2.5	2.3	2.0	4.2	4.2	4.1	4.0	3.9	3.8	3.6	3.4	3.2	3.0	2.7	2.4
		−1.28	.10	3.1	3.0	3.0	2.9	2.8	2.6	2.5	2.3	2.1	1.8	1.6	1.3	3.5	3.5	3.4	3.3	3.2	3.1	2.9	2.7	2.5	2.3	2.0	1.7
		−1.64	.05	2.5	2.5	2.4	2.3	2.2	2.1	1.9	1.7	1.5	1.3	1.0	0.7	2.9	2.9	2.8	2.7	2.6	2.5	2.3	2.1	1.9	1.7	1.4	1.1
	LE average	1.64	.95	8.7	8.6	8.4	8.2	8.0	7.8	7.5	7.3	7.0	6.6	6.3	5.9	9.1	9.0	8.8	8.7	8.5	8.2	8.0	7.7	7.4	7.1	6.7	6.3
		1.28	.90	8.1	8.0	7.8	7.7	7.5	7.2	7.0	6.7	6.4	6.1	5.7	5.3	8.5	8.4	8.3	8.1	7.9	7.7	7.4	7.1	6.8	6.5	6.1	5.8
		.84	.80	7.4	7.3	7.1	7.0	6.8	6.5	6.3	6.0	5.7	5.4	5.0	4.6	7.8	7.7	7.6	7.4	7.2	7.0	6.7	6.4	6.1	5.8	5.4	5.1
		0	.50	6.1	6.0	5.8	5.6	5.4	5.2	4.9	4.7	4.4	4.0	3.7	3.3	6.5	6.4	6.2	6.1	5.9	5.6	5.4	5.1	4.8	4.5	4.1	3.7
		−.84	.20	4.8	4.6	4.5	4.3	4.1	3.9	3.6	3.3	3.0	2.7	2.4	2.0	5.2	5.1	4.9	4.7	4.5	4.3	4.0	3.8	3.5	3.1	2.8	2.4
		−1.28	.10	4.1	3.9	3.8	3.6	3.4	3.2	2.9	2.6	2.3	2.0	1.7	1.3	4.5	4.4	4.2	4.0	3.8	3.6	3.3	3.1	2.8	2.4	2.1	1.7
		−1.64	.05	3.5	3.4	3.2	3.0	2.8	2.6	2.4	2.1	1.8	1.4	1.1	0.7	3.9	3.8	3.6	3.5	3.3	3.0	2.8	2.5	2.2	1.9	1.5	1.1
	LE high	1.64	.95	9.3	9.1	8.9	8.7	8.5	8.2	7.9	7.6	7.3	6.9	6.5	6.1	9.7	9.5	9.4	9.1	8.9	8.6	8.4	8.0	7.7	7.3	6.9	6.5
		1.28	.90	8.7	8.6	8.4	8.2	7.9	7.6	7.4	7.0	6.7	6.3	5.9	5.5	9.1	9.0	8.8	8.6	8.3	8.1	7.8	7.5	7.1	6.8	6.4	6.0
		.84	.80	8.0	7.9	7.7	7.5	7.2	6.9	6.7	6.3	6.0	5.6	5.3	4.8	8.4	8.3	8.1	7.9	7.6	7.4	7.1	6.8	6.4	6.1	5.7	5.3
		0	.50	6.7	6.5	6.3	6.1	5.9	5.6	5.3	5.0	4.7	4.3	3.9	3.5	7.1	7.0	6.8	6.5	6.3	6.0	5.8	5.4	5.1	4.7	4.3	3.9
		−.84	.20	5.4	5.2	5.0	4.8	4.6	4.3	4.0	3.7	3.3	3.0	2.6	2.2	5.8	5.6	5.4	5.2	5.0	4.7	4.4	4.1	3.8	3.4	3.0	2.6
		−1.28	.10	4.7	4.5	4.3	4.1	3.9	3.6	3.3	3.0	2.6	2.3	1.9	1.5	5.1	4.9	4.7	4.5	4.3	4.0	3.7	3.4	3.1	2.7	2.3	1.9
		−1.64	.05	4.1	3.9	3.7	3.5	3.3	3.0	2.7	2.4	2.1	1.7	1.3	0.9	4.5	4.4	4.2	4.0	3.7	3.4	3.2	2.8	2.5	2.1	1.7	1.3
Auditory VLT	LE low	1.64	.95	8.1	8.0	8.0	7.9	7.8	7.6	7.5	7.3	7.1	6.8	6.6	6.3	8.5	8.4	8.4	8.3	8.2	8.0	7.9	7.7	7.5	7.2	7.0	6.7
		1.28	.90	7.5	7.4	7.4	7.3	7.2	7.0	6.9	6.7	6.5	6.2	6.0	5.7	7.9	7.9	7.8	7.7	7.6	7.5	7.3	7.1	6.9	6.7	6.4	6.1
		.84	.80	6.8	6.7	6.7	6.6	6.5	6.3	6.2	6.0	5.8	5.5	5.3	5.0	7.2	7.2	7.1	7.0	6.9	6.8	6.6	6.4	6.2	6.0	5.7	5.4
		0	.50	5.5	5.4	5.4	5.3	5.2	5.0	4.9	4.7	4.5	4.2	4.0	3.7	5.9	5.8	5.8	5.7	5.6	5.4	5.3	5.1	4.9	4.6	4.4	4.1
		−.84	.20	4.1	4.1	4.0	3.9	3.8	3.7	3.5	3.3	3.1	2.9	2.6	2.3	4.5	4.5	4.4	4.4	4.2	4.1	3.9	3.8	3.5	3.3	3.0	2.8
		−1.28	.10	3.4	3.4	3.3	3.2	3.1	3.0	2.8	2.6	2.4	2.2	1.9	1.6	3.9	3.8	3.8	3.7	3.6	3.4	3.3	3.1	2.9	2.6	2.4	2.1
		−1.64	.05	2.9	2.8	2.8	2.7	2.6	2.4	2.3	2.1	1.9	1.6	1.4	1.1	3.3	3.2	3.2	3.1	3.0	2.8	2.7	2.5	2.3	2.0	1.8	1.5
	LE average	1.64	.95	9.0	8.9	8.8	8.6	8.4	8.2	7.9	7.6	7.3	7.0	6.6	6.3	9.5	9.3	9.2	9.0	8.8	8.6	8.3	8.0	7.7	7.4	7.1	6.7
		1.28	.90	8.5	8.3	8.2	8.0	7.8	7.6	7.3	7.1	6.8	6.4	6.1	5.7	8.9	8.8	8.6	8.4	8.2	8.0	7.8	7.5	7.2	6.8	6.5	6.1
		.84	.80	7.8	7.6	7.5	7.3	7.1	6.9	6.6	6.4	6.1	5.7	5.4	5.0	8.2	8.1	7.9	7.7	7.5	7.3	7.1	6.8	6.5	6.2	5.8	5.4
		0	.50	6.4	6.3	6.2	6.0	5.8	5.6	5.3	5.0	4.7	4.4	4.0	3.7	6.9	6.7	6.6	6.4	6.2	6.0	5.7	5.5	5.1	4.8	4.5	4.1
		−.84	.20	5.1	5.0	4.8	4.7	4.5	4.2	4.0	3.7	3.4	3.1	2.7	2.3	5.5	5.4	5.3	5.1	4.9	4.6	4.4	4.1	3.8	3.5	3.1	2.8
		−1.28	.10	4.4	4.3	4.1	4.0	3.8	3.5	3.3	3.0	2.7	2.4	2.0	1.6	4.8	4.7	4.6	4.4	4.2	4.0	3.7	3.4	3.1	2.8	2.4	2.1
		−1.64	.05	3.8	3.7	3.6	3.4	3.2	3.0	2.7	2.4	2.1	1.8	1.4	1.1	4.3	4.1	4.0	3.8	3.6	3.4	3.1	2.9	2.5	2.2	1.9	1.5
	LE high	1.64	.95	9.6	9.5	9.3	9.1	8.8	8.6	8.3	8.0	7.6	7.3	6.9	6.5	10.1	9.9	9.7	9.5	9.3	9.0	8.7	8.4	8.1	7.7	7.3	6.9
		1.28	.90	9.1	8.9	8.7	8.5	8.3	8.0	7.7	7.4	7.1	6.7	6.3	5.9	9.5	9.3	9.1	8.9	8.7	8.4	8.1	7.8	7.5	7.1	6.7	6.3
		.84	.80	8.4	8.2	8.0	7.8	7.6	7.3	7.0	6.7	6.4	6.0	5.6	5.2	8.8	8.6	8.4	8.2	8.0	7.7	7.4	7.1	6.8	6.4	6.0	5.6
		0	.50	7.0	6.9	6.7	6.5	6.2	6.0	5.7	5.4	5.0	4.7	4.3	3.9	7.5	7.3	7.1	6.9	6.7	6.4	6.1	5.8	5.5	5.1	4.7	4.3
		−.84	.20	5.7	5.5	5.4	5.1	4.9	4.6	4.4	4.0	3.7	3.3	2.9	2.5	6.1	6.0	5.8	5.6	5.3	5.1	4.8	4.5	4.1	3.8	3.4	3.0
		−1.28	.10	5.0	4.9	4.7	4.4	4.2	3.9	3.7	3.3	3.0	2.6	2.2	1.8	5.4	5.3	5.1	4.9	4.6	4.4	4.1	3.8	3.4	3.1	2.7	2.3
		−1.64	.05	4.4	4.3	4.1	3.9	3.6	3.4	3.1	2.8	2.4	2.1	1.7	1.3	4.9	4.7	4.5	4.3	4.1	3.8	3.5	3.2	2.9	2.5	2.1	1.7

Note. LE = Level of education; Cum. prob. = cumulative probability.

Table A2. Normative VLT data for the Delta score stratified by Mode of presentation, Sex, and, Age (25, 30, . . . , 80 years). The raw test score leading to a particular Z-value is given for Z-values indicating the percentiles 5, 10, 20, 50, 80, 90, and, 95

	Z value	Cum. prob.	Male Age in years 25	30	35	40	45	50	55	60	65	70	75	80	Female Age in years 25	30	35	40	45	50	55	60	65	70	75	80
Visual VLT / LE low/average/high	1.64	.95	8.0	7.9	7.9	7.8	7.7	7.7	7.6	7.5	7.5	7.4	7.3	7.3	8.5	8.4	8.4	8.3	8.2	8.2	8.1	8.0	8.0	7.9	7.8	7.8
	1.28	.90	7.3	7.3	7.2	7.1	7.1	7.0	6.9	6.9	6.8	6.7	6.7	6.6	7.8	7.8	7.7	7.6	7.6	7.5	7.4	7.4	7.3	7.2	7.2	7.1
	.84	.80	6.5	6.4	6.4	6.3	6.2	6.2	6.1	6.0	6.0	5.9	5.8	5.8	7.0	6.9	6.9	6.8	6.7	6.7	6.6	6.5	6.5	6.4	6.3	6.3
	0	.50	4.9	4.9	4.8	4.7	4.7	4.6	4.5	4.5	4.4	4.3	4.3	4.2	5.4	5.4	5.3	5.2	5.2	5.1	5.0	5.0	4.9	4.8	4.8	4.7
	−.84	.20	3.3	3.3	3.2	3.2	3.1	3.0	3.0	2.9	2.8	2.8	2.7	2.6	3.9	3.8	3.7	3.7	3.6	3.5	3.5	3.4	3.3	3.3	3.2	3.1
	−1.28	.10	2.5	2.5	2.4	2.3	2.3	2.2	2.1	2.1	2.0	1.9	1.9	1.8	3.0	3.0	2.9	2.8	2.8	2.7	2.6	2.6	2.5	2.4	2.4	2.3
	−1.64	.05	1.9	1.8	1.7	1.7	1.6	1.5	1.5	1.4	1.3	1.3	1.2	1.1	2.4	2.3	2.2	2.2	2.1	2.0	2.0	1.9	1.8	1.8	1.7	1.6
Auditory VLT / LE low/average/high	1.64	.95	7.1	7.1	7.0	6.9	6.9	6.8	6.7	6.7	6.6	6.6	6.5	6.4	7.6	7.6	7.5	7.4	7.4	7.3	7.2	7.2	7.1	7.1	7.0	6.9
	1.28	.90	6.5	6.4	6.3	6.3	6.2	6.1	6.1	6.0	5.9	5.9	5.8	5.7	7.0	6.9	6.8	6.8	6.7	6.6	6.6	6.5	6.4	6.4	6.3	6.2
	.84	.80	5.6	5.6	5.5	5.4	5.4	5.3	5.2	5.2	5.1	5.1	5.0	4.9	6.1	6.1	6.0	5.9	5.9	5.8	5.8	5.7	5.6	5.6	5.5	5.4
	0	.50	4.1	4.0	3.9	3.9	3.8	3.7	3.7	3.6	3.5	3.5	3.4	3.4	4.6	4.5	4.4	4.4	4.3	4.2	4.2	4.1	4.0	4.0	3.9	3.9
	−.84	.20	2.5	2.4	2.4	2.3	2.2	2.2	2.1	2.0	2.0	1.9	1.8	1.8	3.0	2.9	2.9	2.8	2.7	2.7	2.6	2.5	2.5	2.4	2.3	2.3
	−1.28	.10	1.7	1.6	1.5	1.5	1.4	1.3	1.3	1.2	1.1	1.1	1.0	1.0	2.2	2.1	2.0	2.0	1.9	1.8	1.8	1.7	1.6	1.6	1.5	1.5
	−1.64	.05	1.0	0.9	0.9	0.8	0.7	0.7	0.6	0.5	0.5	0.4	0.3	0.3	1.5	1.4	1.4	1.3	1.2	1.2	1.1	1.0	1.0	0.9	0.8	0.8

Note. LE = Level of education; Cum. prob. = cumulative probability.

Table A3. Normative VLT data for the Delayed recall score stratified by Level of education, Sex, and, Age (25, 30, . . . , 80 years). The raw test score leading to a particular Z-value is given for Z-values indicating the percentiles 5, 10, 20, 50, 80, 90, and, 95

	Z value	Cum. prob.	Male Age in years 25	30	35	40	45	50	55	60	65	70	75	80	Female Age in years 25	30	35	40	45	50	55	60	65	70	75	80
Auditory & Visual VLT / LE low	1.64	.95	14.2	14.1	13.9	13.6	13.3	13.0	12.6	12.2	11.7	11.2	10.6	10.0	15.0	15.0	15.0	14.8	14.5	14.2	13.8	13.4	12.9	12.4	11.8	11.2
	1.28	.90	13.3	13.2	13.0	12.7	12.4	12.1	11.7	11.3	10.8	10.3	9.7	9.1	14.5	14.4	14.2	13.9	13.6	13.3	12.9	12.5	12.0	11.5	10.9	10.3
	.84	.80	12.2	12.1	11.9	11.6	11.3	11.0	10.6	10.2	9.7	9.2	8.6	8.0	13.4	13.3	13.1	12.8	12.5	12.2	11.8	11.4	10.9	10.4	9.8	9.2
	0	.50	10.1	10.0	9.8	9.5	9.2	8.9	8.5	8.1	7.6	7.1	6.5	5.9	11.3	11.2	11.0	10.7	10.4	10.1	9.7	9.3	8.8	8.3	7.7	7.1
	−.84	.20	8.0	7.9	7.7	7.4	7.1	6.8	6.4	6.0	5.5	5.0	4.4	3.8	9.2	9.1	8.9	8.6	8.3	8.0	7.6	7.2	6.7	6.2	5.6	5.0
	−1.28	.10	7.0	6.8	6.6	6.3	6.0	5.7	5.3	4.9	4.4	3.9	3.3	2.7	8.1	8.0	7.8	7.5	7.2	6.9	6.5	6.1	5.6	5.1	4.5	3.9
	−1.64	.05	6.1	5.9	5.7	5.4	5.1	4.8	4.4	4.0	3.5	3.0	2.4	1.8	7.2	7.1	6.9	6.6	6.3	6.0	5.6	5.2	4.7	4.2	3.6	3.0
LE average	1.64	.95	15.0	14.9	14.7	14.5	14.2	13.8	13.4	13.0	12.5	12.0	11.4	10.8	15.0	15.0	15.0	15.0	15.0	15.0	14.6	14.2	13.7	13.2	12.6	12.0
	1.28	.90	14.2	14.0	13.8	13.6	13.3	12.9	12.5	12.1	11.6	11.1	10.5	9.9	15.0	15.0	15.0	14.8	14.5	14.1	13.7	13.3	12.8	12.3	11.7	11.1
	.84	.80	13.1	12.9	12.7	12.5	12.2	11.8	11.4	11.0	10.5	10.0	9.4	8.8	14.3	14.1	13.9	13.7	13.4	13.0	12.6	12.2	11.7	11.2	10.6	10.0
	0	.50	11.0	10.8	10.6	10.4	10.1	9.7	9.3	8.9	8.4	7.9	7.3	6.7	12.2	12.0	11.8	11.6	11.3	10.9	10.5	10.1	9.6	9.1	8.5	7.9
	−.84	.20	8.9	8.7	8.5	8.3	8.0	7.6	7.2	6.8	6.3	5.8	5.2	4.6	10.1	9.9	9.7	9.5	9.2	8.8	8.4	8.0	7.5	7.0	6.4	5.8
	−1.28	.10	7.8	7.6	7.4	7.2	6.9	6.5	6.1	5.7	5.2	4.7	4.1	3.5	9.0	8.8	8.6	8.4	8.1	7.7	7.3	6.9	6.4	5.9	5.3	4.7
	−1.64	.05	6.9	6.7	6.5	6.3	6.0	5.6	5.2	4.8	4.3	3.8	3.2	2.6	8.1	7.9	7.7	7.5	7.2	6.8	6.4	6.0	5.5	5.0	4.4	3.8
LE high	1.64	.95	15.0	15.0	15.0	14.9	14.6	14.2	13.9	13.4	12.9	12.4	11.9	11.2	15.0	15.0	15.0	15.0	15.0	15.0	15.0	14.6	14.1	13.6	13.1	12.4
	1.28	.90	14.6	14.4	14.2	14.0	13.7	13.3	13.0	12.5	12.0	11.5	11.0	10.3	15.0	15.0	15.0	15.0	14.9	14.5	14.2	13.7	13.2	12.7	12.2	11.5
	.84	.80	13.5	13.3	13.1	12.9	12.6	12.2	11.9	11.4	11.0	10.4	9.9	9.2	14.7	14.5	14.3	14.1	13.8	13.4	13.1	12.6	12.1	11.6	11.1	10.4
	0	.50	11.4	11.3	11.0	10.8	10.5	10.2	9.8	9.3	8.9	8.3	7.8	7.2	12.6	12.4	12.2	12.0	11.7	11.3	11.0	10.5	10.1	9.5	9.0	8.3
	−.84	.20	9.3	9.2	8.9	8.7	8.4	8.1	7.7	7.2	6.8	6.2	5.7	5.1	10.5	10.4	10.1	9.9	9.6	9.3	8.9	8.4	8.0	7.4	6.9	6.3
	−1.28	.10	8.2	8.1	7.8	7.6	7.3	7.0	6.6	6.1	5.7	5.1	4.6	4.0	9.4	9.3	9.0	8.8	8.5	8.2	7.8	7.3	6.9	6.3	5.8	5.2
	−1.64	.05	7.3	7.2	7.0	6.7	6.4	6.1	5.7	5.2	4.8	4.2	3.7	3.1	8.5	8.4	8.1	7.9	7.6	7.3	6.9	6.4	6.0	5.4	4.9	4.3

Note. LE = Level of education; Cum. prob. = cumulative probability.

Table A4. Normative VLT data for the Recognition score stratified by Sex and Age (25, 30, . . . , 80 years). The raw test score leading to a particular Z-value is given for Z-values indicating the percentiles 5, 10, 20, 50, 80, 90, and, 95

	Z value	Cum. prob.	Male — Age in years												Female — Age in years											
			25	30	35	40	45	50	55	60	65	70	75	80	25	30	35	40	45	50	55	60	65	70	75	80
Auditory & Visual VLT / LE low/average/high	1.64	.95	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0
	1.28	.90	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0
	.84	.80	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	14.9	14.5	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	14.9
	0	.50	14.1	14.2	14.1	14.1	13.9	13.8	13.6	13.4	13.1	12.8	12.4	12.1	14.5	14.6	14.5	14.5	14.3	14.2	14.0	13.8	13.5	13.2	12.8	12.5
	−.84	.20	12.8	12.9	12.8	12.8	12.6	11.9	11.7	11.5	10.6	10.3	9.9	9.6	13.4	13.5	13.4	13.4	13.2	13.1	12.7	11.9	11.6	10.7	10.3	10.0
	−1.28	.10	12.2	12.2	12.1	12.1	12.0	10.9	10.7	10.5	9.3	9.0	8.6	8.2	12.9	12.9	12.9	12.8	12.7	12.5	12.0	10.9	10.6	9.4	9.0	8.6
	−1.64	.05	11.6	11.6	11.6	11.5	11.4	10.1	9.9	9.7	8.2	7.9	7.6	7.2	12.4	12.4	12.4	12.3	12.2	12.1	11.5	10.1	9.8	8.3	8.0	7.6

Note. LE = Level of education; Cum. prob. = cumulative probability.