OXFORD UNIVERSITY PRESS

Archives of CLINICAL NEUROPSYCHOLOGY

Archives of Clinical Neuropsychology 27 (2012) 510–519

Comparison of Verbal Episodic Memory Measures: Consortium to Establish a Registry for Alzheimer's Disease—Neuropsychological Assessment Battery (CERAD-NAB) versus California Verbal Learning Test (CVLT)

Irene R. Beck[1†], Alexa Gagneux-Zurbriggen[1†], Manfred Berres[2], Kirsten I. Taylor[1,3], Andreas U. Monsch[1,*]

[1]*Memory Clinic, Department of Geriatrics, University Hospital Basel, Basel, Switzerland*
[2]*Department of Mathematics and Technology, University of Applied Sciences Koblenz, Koblenz, Germany*
[3]*Centre for Speech, Language and the Brain, Department of Experimental Psychology, University of Cambridge, Cambridge, UK*

*Corresponding author at: Memory Clinic, Department of Geriatrics, University Hospital Basel, Schanzenstrasse 55, Basel 4031, Switzerland. Tel.: +41 61 265 31 93; fax: +41 61 265 37 94.
E-mail address: andreas.monsch@unibas.ch (Andreas U. Monsch).

Accepted 23 May 2012

Abstract

Episodic memory is affected early in the course of dementia. Two well-established tests to assess verbal episodic memory functioning are the Word List task from the Consortium to Establish a Registry for Alzheimer's disease Neuropsychological Assessment Battery (CERAD-NAB) and the California Verbal Learning Test (CVLT). In clinical and/or research settings, patients are typically administered either one or the other test, making statistical comparisons difficult. This study aimed to (i) compare the z-scores of these two tests in patients with MCI and different types of dementia and (ii) establish formulae to transform CERAD-NAB scores into CVLT scores and vice versa. Sixty-five patients completed both tests for the first time and within 10 days of each other. Pearson correlation coefficients indicated that the two tests assess similar aspects of episodic memory and that the CVLT is more sensitive to subtle episodic memory impairments. Finally, conversion formulae are provided and their implementation illustrated.

Keywords: Alzheimer's disease; Dementia; Elderly/geriatrics/aging; Learning and memory; Mild cognitive impairment; Statistical methods

Introduction

Episodic memory impairments are among the most common early cognitive signs of dementia (Bäckman, Small, & Fratiglioni, 2001; Cullum, Filley, & Kozora, 1995; Fox, Olin, Erblich, Ippen, & Schneider, 1998; Greenaway et al., 2006; Twamley, Legendre Ropacki, & Bondi, 2006; Welsh, Butters, Hughes, Mohs, & Heyman, 1991). Two well-established tests to assess episodic memory functioning in patients with mild cognitive impairment (MCI) or early dementia are the Consortium to Establish a Registry for Alzheimer's disease—Neuropsychological Assessment Battery (CERAD-NAB) Word List task (Morris, Mohs, Rogers, Fillenbaum, & Heyman, 1988) and the California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan, & Ober, 1987). Although both tests assess similar aspects of verbal episodic memory, most notably encoding, recall and recognition, important differences exist regarding test construction and demands. In particular, the CVLT was constructed to assess a broader range of verbal episodic memory-related functions, and is considered more demanding than the CERAD-NAB Word List task. Consequently, patients with progressive neurodegenerative diseases are often

† These authors contributed equally to this work.

doi:10.1093/arclin/acs056 Advance Access publication on 19 June 2012

administered the CVLT in the early stages of their disease and the CERAD-NAB in more advanced stages, such that the statistical longitudinal evaluation of episodic memory functioning become difficult. These task differences are also relevant in research settings when investigators wish to compare the performances of patients examined with either one of these memory tests. Thus, the aims of this study were to examine whether the two verbal learning tests assess similar aspects of episodic memory and whether transformation formulae could be developed to transform the tests' key variables into one another.

Patients with amnestic MCI as well as those with different forms of dementia frequently complain about their episodic memory deficits, such as forgetting names or losing things (Jonker, Launer, Hooijer, & Lindeboom, 1996). Indeed, neuropsychological studies have demonstrated that episodic learning and recall best discriminate between healthy controls and preclinical Alzheimer's disease (AD) patients and between preclinical AD and probable AD patients (Collie & Maruff, 2000; Twamley et al., 2006). Depending on the type of dementia, patients show limited learning, rapid forgetting, and/or impaired recognition of task information (Bäckman et al., 2001; Cullum et al., 1995; Fox et al., 1998; Greenaway et al., 2006; Grober, Lipton, Hall, & Crystal, 2000; Romàn, Erkinjuntti, Wallin, Pantoni, & Chui, 2002; Welsh et al., 1991). Consistent with these findings, magnetic resonance imaging studies typically report a significant reduction in hippocampal and entorhinal cortex volumes in patients with MCI and AD (De Leon et al., 1997; Du et al., 2001; Fox et al., 1996; Krasuski et al., 1998), two structures essential for episodic memory (Rempel-Clower, Zola, Squire, & Amaral, 1996; Squire & Zola-Morgan, 1991; Victor & Agamanolis, 1990). Given the importance of episodic memory deficits in the diagnosis of dementia, the assessment of verbal episodic memory is a crucial part of the neuropsychological examination of patients with suspected MCI or early dementia.

The standardized neuropsychological assessment of dementia includes tests which measure the ability to learn new information, e.g., a list of words over a number of trials. The sum of immediately recalled ("immediate recall") words constitutes "encoding", i.e., the storage of incoming information into a mental representation (Butters, Delis, & Lucas, 1995). After a delay ranging between roughly 5 and 20 min, patients are requested to recall the encoded words, providing a measure of "delayed recall". Finally, many standardized assessments of verbal episodic memory end with a recognition test.

One well-established test of verbal episodic memory in dementia is the Word List task from the CERAD-NAB (Morris et al., 1988), a standardized serial learning task. Another widely used test is the CVLT (Delis et al., 1987), which assesses the same memory components as the CERAD-NAB Word List task, and additionally quantifies the different strategies people use to learn verbal information. Given the popularity of these two tests, it is important to know their relative usefulness and their possible interchangeability. As noted above, both tests assess similar aspects of verbal episodic learning and memory: encoding, delayed recall, and recognition. However, there are substantial differences in the task demands and format. For example, compared with the CERAD-NAB Word List task, the CVLT has more items to be learned (16 vs. 10) and longer words on average (2.37 vs. 1.70 syllables). Furthermore, the CVLT learning list has more learning trials (5 vs. 3), a longer delay between the last learning trial and the delayed recall (about 20 min vs. about 5 min) is presented in only one modality (auditory-verbal vs. auditory-verbal and orthographic) and has more distractor items on the recognition list (64% vs. 50%). On the other hand, items on the CVLT learning list are always presented in the same order on each trial, whereas the order of item presentation varies across trials on the CERAD-NAB such that patients are not able to benefit from within-list primacy and recency effects. Moreover, the items on the CVLT list can be categorized into semantic categories, which may facilitate performance, whereas the words on the CERAD-NAB Word List are not obviously semantically related. This latter point is of special interest for patients with dementia since a number of studies have indicated that semantic memory functioning is affected in AD, potentially in the earliest stages of the disease (Chen et al., 2001; Elias et al., 2000; Martin & Fedio, 1983; Masur, Sliwinski, Lipton, Blau, & Crystal, 1994; Monsch et al., 1992; Nebes, 1989; Salmon, Heindel, & Lange, 1999; Saxton et al., 2004). For example, Weingartner and colleagues (1981) found that the use of semantic categories in a word list task improved the performance in healthy controls but not that of patients with dementia. Thus, the performance on the two memory tests will be affected by differential task demands, with the CVLT placing greater cognitive demands on participants, and taxing other cognitive processes beyond pure verbal episodic memory functioning.

The difference in the difficulty of the two tests has major implications for clinical practice. In the case of a suspected MCI or early dementia patient, clinicians generally choose the most demanding episodic memory task to detect even slight memory impairments. However, low-functioning or more advanced patients may be overwhelmed with the CVLT's 16-item word list presented over five trials. In such cases, a step-down battery such as the CERAD-NAB Word List task with 10 words and three trials may be more appropriate. Moreover, it may be desirable to compare different patients' episodic memory functioning assessed with different tests, either in the clinical or in the research setting. Also the evaluation of clinical progression necessitates the consecutive assessment of verbal episodic memory performance, which may require different tests over time. In these latter instances, it would be helpful to compare the performance on the more difficult CVLT with that of the less-demanding CERAD-NAB Word List task, i.e., to know the relationship between key variables of the two tests. The purpose of this study

was therefore to (i) compare *z*-scores of key CERAD-NAB word list with the CVLT variables in patients with MCI and different types of early-stage dementia and (ii) establish formulae to transform CERAD-NAB *z*-scores into CVLT *z*-scores and vice versa for use in clinical practice and research. We hypothesized that the performance on the two word list tasks will be significantly correlated, reflecting the measurement of similar aspects of verbal episodic memory functioning. Further, we hypothesized that the greater cognitive demands made by the CVLT will be manifested in poorer patient performance on this task compared with the CERAD-NAB Word List test.

Method

Participants included 65 patients (27 women and 38 men) examined at the Memory Clinic, Department of Geriatrics, University Hospital Basel, Switzerland, between 1998 and 2006. The sample consisted of 29 patients with a diagnosis of MCI according to the Winblad and colleagues (2004) criteria, 33 patients with a diagnosis of probable AD according to the criteria outlined by the National Institute for Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA; McKhann et al., 1984) and DSM-IV (American Psychiatric Association, 1994) criteria for AD, two with a diagnosis of probable vascular dementia according to the criteria outlined by the National Institute of Neurological Disorders and Stroke and the Association Internationale pour la Recherche et l'Enseignement en Neurosciences (NINDS-AIREN; Romàn et al., 1993), and one with frontotemporal lobar degeneration according to the Neary and colleagues (1998) criteria. The mean age was 69.3 years ($SD = 10.1$; range = 50–86) and the mean educational level was 12.6 years ($SD = 3.4$; range = 7–20). Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) scores ranged between 22 and 30, with a mean of 27.2 ($SD = 2.3$). German versions of the CERAD-NAB and the CVLT were administered as part of a comprehensive standardized neuropsychological assessment of dementia. All patients completed the CERAD-NAB and the CVLT for the first time and within 10 days of each other with a mean time difference of 2.00 days ($SD = 2.91$, range = 0–10). Twenty-nine were participants of the BASEL project (BAsel Study on the ELderly; Monsch et al., 2000), a longitudinal study assessing the cognitive performance in older individuals. Twenty of these patients first completed the CERAD-NAB Word List task and afterwards the CVLT, while nine completed the tests in the reverse order. The remaining 36 patients were examined as part of the clinical routine at the Memory Clinic, where they were first administered the CVLT and then the CERAD-NAB Word List task, both on the same day. The retrospective data analysis was approved by the Ethics Committee of both Basels, Switzerland. All patients spoke German as their native language.

The CERAD-NAB Word List task presents patients with 10 unrelated words which are printed on 10 different cards. The patient is instructed to read each word out loud and to remember each word. After the presentation of all 10 words, immediate recall is assessed. All words are presented twice again, each time in a different order, and the patient is asked to recall as many words as possible following each trial. The maximum score of correct responses (encoding) is 30. After about 5 min of nonverbal testing, e.g., copying geometric figures, the patient is required to recall the 10 words again (delayed recall). Immediately after the delayed recall, a recognition task is administered, during which the patient is presented with 20 words, 10 words from the original word list and 10 new (distractor) items, and asked to identify the original words.

The CVLT contains two different word lists with 16 shopping items each ("Monday list" and "Tuesday list") from four semantic categories (tools, articles of clothing, drinks, and fruits). In the encoding phase, the 16 shopping items from the Monday list are read aloud to the patient over five trials, each time in the same order. After each presentation, patients are instructed to recall as many items as possible. The maximum score of correct responses is 80 (encoding). Following the immediate recall of Trial 5, patients are presented with the interference Tuesday list and required to recall as many items as possible from this list. Patients are then again asked to recall the words from the Monday list, and are subsequently cued with each of the four category labels. After a 20-min delay with nonverbal testing, free and cued recall tasks of the Monday list are administered. Finally, a recognition list with 44 shopping items (16 from the Monday list and 28 distractor items) is read to the patient out loud and he/she decides if the word was on the Monday list or not.

Based on the methodology described by Berres, Monsch, Bernasconi, Thalmann, and Stähelin (2000; see also Berres, Zehnder, Bläsi, & Monsch, 2008), all raw scores were transformed into demographically adjusted *z*-scores. *z*-Scores of key variables representing (i) encoding, (ii) recall, and (iii) recognition performance on the two word lists were compared using Pearson correlation coefficients with a 0.05 alpha level. Paired *t*-tests were performed to test for differences between the key variables, with alpha set to 0.05. Linear regression models are usually used to transform one test-score y into another test-score x. Since we aimed to predict CVLT scores (y) from CERAD-NAB scores (x) and vice versa, two linear regression models could have been used to produce two different regression lines. The variance of predictions from ordinary least square (OLS) linear regression is smaller than the variance of the observed values, except if $r^2 = 1$. This property of OLS

regression is undesirable for the present purposes. An alternative method is to calculate the first principal component (PC), which represents a compromise between the two regression models (Berres & Volland, 2011; Volland, 2011). Contrary to linear regressions, PCs remain the same if the roles of x and y are switched. Thus, to describe the linear relationships between the z-scores of each pair of key variables, PC analyses were performed on the covariance matrices of the two corresponding sets of z-scores, and the first component was used to compute the linear transformations between the two test variables. Thus, the following equation transformed the z-scores from a CERAD-NAB Word List variable (x) into a z-score from CVLT variable (y)

$$y = \bar{y} + \frac{v_2}{v_1}(x - \bar{x}), \tag{1}$$

where v_1 and v_2 are the coordinates of the first eigenvector. The variance of the prediction of y given x is

$$s^2_{y|x} = \frac{1}{n}s^2_y(1 - r^2) + \frac{d_2 d_1}{(n-1)(d_2 - d_1)^2}\frac{1}{v_1^4}(x - \bar{x})^2, \tag{2}$$

where d_1 and d_2 are eigenvalues of the covariance matrix, $s^2{}_y$ is the variance of y, and n is the sample size and the factor preceding $(x - \bar{x})^2$ is the variance of v_2/v_1 (cf. Mardia, Kent, & Bibby, 1979). The approximate 95% confidence band for the line of the first PC is

$$\bar{y} + \frac{v_2}{v_1}(x - \bar{x}) \pm 1.96\, s_{y|x}. \tag{3}$$

Solving for x in (1) gives the equation

$$x = \bar{x} + \frac{v_1}{v_2}(y - \bar{y}), \tag{4}$$

to transform z-scores from CVLT (y) variables into CERAD-NAB Word List z-scores (x). The variance of the prediction of x from y is calculated as follows:

$$s^2_{x|y} = \frac{s^2_x}{n}(1 - r^2) + \frac{d_2 d_1}{(n-1)(d_2 - d_1)^2}\frac{1}{v_2^4}(y - \bar{y})^2. \tag{5}$$

Results

The means, standard deviations, and ranges of the z-scores of the selected CERAD-NAB and CVLT variables are presented in Table 1. As the mean z-scores indicate, patients demonstrated limited learning of verbal information over serial trials, rapid forgetting after some delay, and impaired recognition on both the CERAD-NAB and the CVLT word lists. The correlations between the selected CVLT and CERAD-NAB z-scores were significant for all key variables (encoding: $r = .727$, $p < .01$; recall: $r = .762$, $p < .01$; recognition: $r = .554$, $p < .01$), indicating that the two tests measured similar aspects of verbal episodic memory functioning. To determine whether the z-scores of the key variables were significantly different from each other, we performed paired t-tests. P-values of these t-tests (see Table 1) revealed that patients performed significantly worse on the CVLT compared with the CERAD-NAB Word List task (encoding: $t(63) = 3.95$; $p < .001$; recall: $t(63) = 6.50$; $p < .001$;

Table 1. Descriptive statistics of key variables (z-scores) from the CERAD-NAB and CVLT

	CERAD-NAB		CVLT		p-Value
	Mean (*SD*)	Range	Mean (*SD*)	Range	
Encoding	−1.28 (1.42)	−5.13 to 1.43	−1.80 (1.40)	−4.54 to 1.99	.001
Recall	−1.27 (1.30)	−4.07 to 2.01	−2.07 (1.55)	−4.91 to 1.80	.001
Recognition	−0.92 (1.34)	−3.93 to 1.21	−1.59 (1.83)	−4.61 to 1.41	.001

Note: CERAD-NAB = Consortium to Establish a Registry for Alzheimer's Disease—Neuropsychological Assessment Battery (Morris et al., 1988); CVLT = California Verbal Learning Test (Delis et al., 1987).

Table 2. Eigenvectors and eigenvalues corresponding to the first PC

	First eigenvector		Eigenvalues	
	v_1	v_2	d_1	d_2
Encoding	−0.715	0.699	3.419	0.541
Recall	0.623	−0.782	3.648	0.474
Recognition	0.496	−0.869	4.108	0.991

Note: PC = Principal Component.

Table 3. Formulae to transform performance (*z*-scores) on key CERAD-NAB word list task and CVLT variables into one another

	CERAD-NAB	CVLT
Encoding	0.558 + 1.022 × CVLT	−0.546 + 0.978 × CERAD-NAB
Recall	0.402 + 0.797 × CVLT	−0.505 + 1.255 × CERAD-NAB
Recognition	0.017 + 0.571 × CVLT	−0.030 + 1.753 × CERAD-NAB

Note: CERAD-NAB = Consortium to Establish a Registry for Alzheimer's Disease—Neuropsychological Assessment Battery (Morris et al., 1988); CVLT = California Verbal Learning Test (Delis et al., 1987).

recognition: $t(63) = 3.61$; $p < .001$), indicating that the CVLT was more difficult for patients with MCI and early dementia than the CERAD-NAB Word List task.

Calculations of the slopes and intercepts from the coefficients of the first eigenvector and the eigenvalues of the PCs (Table 2) led to the linear equations to transform CERAD-NAB *z*-scores into CVLT *z*-scores and vice versa (cf. (1) and (4); see Table 3).

The 95% prediction intervals of the PC lines (cf. (3) and Fig. 1) estimate the accuracy of the CERAD-NAB and CVLT transformation. To estimate the errors of the predicted *z*-scores, we assessed the widths of their 95% confidence intervals. The asymptotic 95% confidence band for the CVLT encoding variable had widths between 0.24 and 0.67, for recall the widths were between 0.26 and 0.77, and for recognition between 0.38 and 2.08, when CERAD-NAB *z*-scores were between −4 and 1. For the CERAD-NAB variables, the asymptotic 95% confidence band had widths between 0.25 and 0.72 for encoding, between 0.21 and 0.56 for recall, and between 0.27 and 0.62 for recognition, when CVLT *z*-scores were between −4 and 1. These confidence bands (see Fig. 1) quantify the amount of inaccuracy that can be expected when predicting the *z*-scores of one word list based on the *z*-scores from the other list.

To explore the clinical utility of these formulae, we created a fictive patient examined four times over 6 years (see Fig. 2). Figure 2A depicts the standard clinical assessment of episodic memory: the CVLT was administered at the first two testing sessions and the CERAD-NAB Word List task was administered as a step-down battery (due to cognitive decline) at follow-ups 2 and 3. As illustrated in Fig. 2A, the course of decline in episodic recall performance is not linear. Counterintuitively, the patient shows better recall performance at follow-up 2 with the CERAD-NAB Word List than at follow-up 1 with the CVLT word list. However, a different course emerged when transformed CVLT recall scores were plotted (see Fig. 2B). Although the course now shows a linear decline in episodic recall, the patient never reaches the critical cut-off point of −1.28, indicating that, overall, the CERAD-NAB is less sensitive to episodic memory impairments. Figure 2C shows the course of episodic recall performance with transformed CERAD-NAB recall scores. The patient reaches a pathological result ($z = -1.43$) in 2001 and an even worse performance in 2003 ($z = -1.91$), demonstrating that the CVLT is a more sensitive tool for detecting episodic memory impairments. A further example demonstrating the clinical utility of the formulae is provided in Fig. 3. These fictive findings demonstrate the benefit of the transformation of CERAD-NAB into CVLT or CVLT into CERAD-NAB key variables in clinical practice: the transformed scores ameliorate differences in task demands such that a more realistic course of memory decline over time becomes apparent. In addition, the bar charts reveal that the CVLT is more sensitive to changes in episodic memory performance than the CERAD-NAB Word List.

Discussion

This study found that MCI and dementia patients attained very similar but not identical results on key variables of the CERAD-NAB Word List task and the CVLT. Specifically, patients obtained significantly lower *z*-scores on the three key CVLT variables compared with the corresponding CERAD-NAB variables. Since situations arise in both clinical and research settings where the groups of patients are administered either one or the other task, or the same patient is administered different

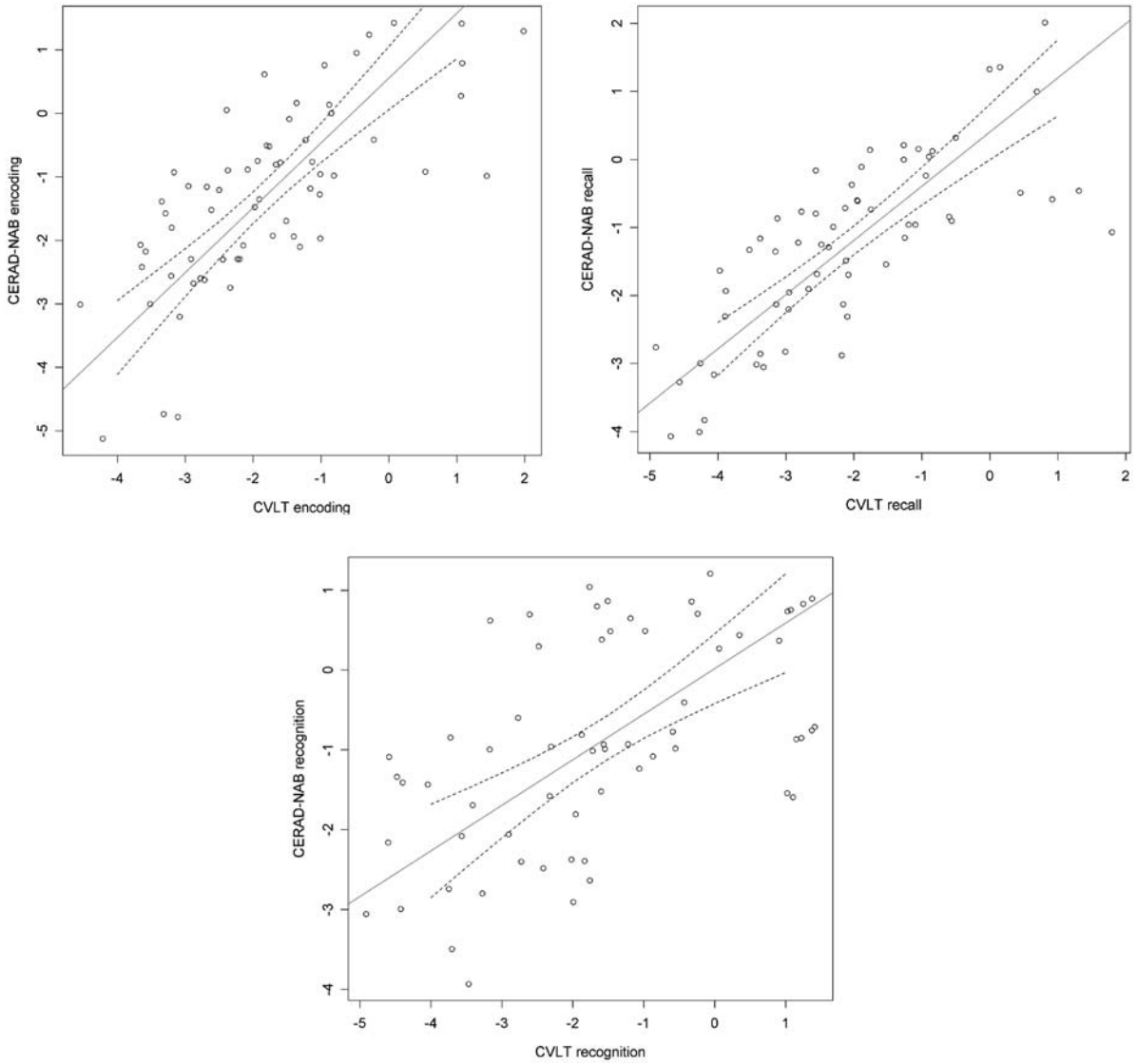

Fig. 1. Scatterplots of *z*-scores from CERAD-NAB Word List task and corresponding CVLT variables with first PCs and 95% confidence bands (dashed).

tasks over time, we developed transformation formulae (Table 3) to convert both tests into one another to facilitate group or intra-individual longitudinal comparisons.

The CVLT and CERAD-NAB verbal episodic memory tasks differ in a number of ways, affecting their respective difficulty and additional cognitive functions engaged during the task. The CVLT is more cognitively demanding, with a greater number of words and learning trials, a single modality of item presentation, a longer delay between the last learning trial and delayed recall, and a larger number of distractor items on recognition. Indeed, we found that patients scored lower on each key CVLT compared with CERAD-NAB variable. This finding is consistent with former studies which compared the CVLT with other commonly used list learning tasks such as the Hopkins Verbal Learning Test (HVLT; Brandt, 1991; Lacritz & Cullum, 1998), the Word List task from the Wechsler Memory Scale-III (WLT; McDowell, Bayless, Moser, Meyers, & Paulsen, 2004; Wechsler, 1997), or the Rey Auditory Verbal Learning Test (RAVLT; Rey, 1964; Stallings, Boake, & Sherer, 1995). Thus, the CVLT appears to be a more sensitive measure of episodic memory deficits than other verbal learning tests. We note that patients scored significantly lower on the CVLT than the CERAD-NAB Word List task despite our comparison of demographically adjusted *z*-scores (i.e., accounting for the effects of age, gender, and educational level) derived from almost identical normative samples (a comparison of the two normative samples is provided at http://www.memoryclinic.ch/).

The formulae we provide allow key variables from the two word list tests to be transformed into one another. Although some limitations apply (see subsequently), these transformations facilitate comparisons of CVLT and CERAD-NAB Word List task

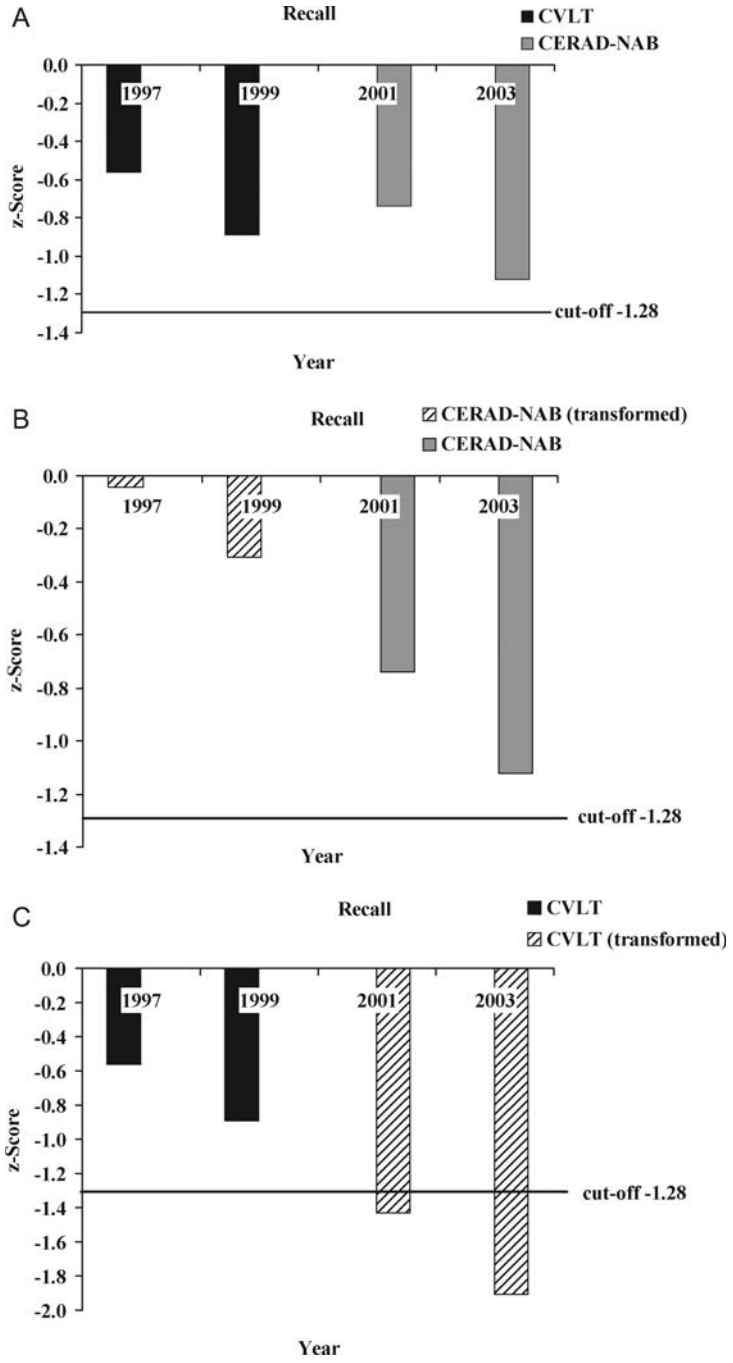

Fig. 2. Recall results of a fictive patient examined four times over 6 years. (A) The patient was administered the CVLT at the first two testing sessions. At follow-ups 2 and 3, he was administered the CERAD-NAB Word List task because his cognitive functioning had declined. (B) Course of recall performance with transformed CVLT recall scores. (C) Course of recall performance with transformed CERAD-NAB Word List task recall scores.

results in clinical and research settings (see Figs. 2 and 3). It must be noted that the width of the approximated 95% confidence bands for the lines of the first PCs indicates that the variances of the predictions are large. Thus, the predicted CERAD-NAB and CVLT *z*-scores should be used and interpreted with caution.

The patients included in the present study were on average well educated, relatively young and mildly cognitive affected. Thus, it is unclear whether the present findings generalize to older or less well-educated patients with more severe dementia. Further, the study had a relatively small sample size, such that it is unclear whether the present results will generalize to the broader population of MCI and AD patients. Moreover, German versions of the two word list tasks were studied here; thus, it

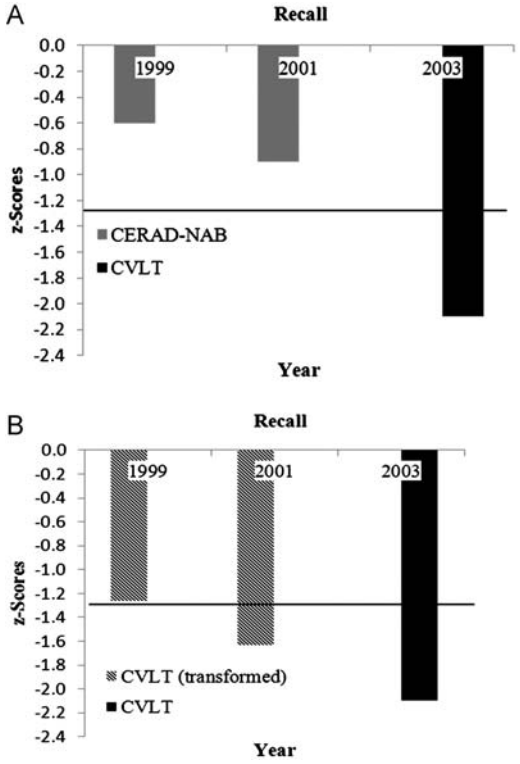

Fig. 3. Recall results of a second fictive patient examined three times over 4 years. (A) The patient was administered the CERAD-NAB Word List task at the first two testing sessions which took place in another hospital. At follow-up 3, the patient was examined at our institution. Since he was well educated, relatively young, and (according to the MMSE) mildly cognitively impaired, he was administered the CVLT. The results show a decline in recall performance over all three testing sessions, with a sudden deterioration at the third testing session. (B) The course of recall performance with transformed CVLT recall scores at baseline and follow-up 1. The transformed scores indicate that the patient would have shown marginal recall performance at baseline and moderate deficits at follow-up 1. The transformed results show a more linear decline in recall performance over the three testing sessions.

remains to be established whether the transformation formulae presented in this study are valid for transforming the performance of English-speaking patients assessed with English versions of the two tasks. We also note that a bias exists in the normative population from which the CVLT *z*-scores were created regarding the ε4 allele, a variant of the apolipoprotein E (ApoE) gene: there were disproportionally more participants with an ApoE ε4 gene in the CVLT normative population than in the unbiased CERAD-NAB normative population. As many studies have shown, the ApoE ε4 gene negatively impacts the cognitive performance, especially episodic memory functions, in cognitively healthy older individuals (Caselli et al., 2004; Deary et al., 2004; Fillenbaum et al., 2001; Small, Rosnick, Fratiglioni, & Bäckman, 2004; Zehnder et al., 2009). Thus, the level of patients' impairments on the CVLT as measured by demographically corrected *z*-scores may have been underestimated. If this were the case, then the CERAD-NAB Word List task and CVLT performance measures would have differed to an even greater extent than reported here. Finally, this study is limited by the investigation of only quantitative aspects of the CVLT and CERAD-NAB Word List tasks. Kaltreider and colleagues (2000) showed that more qualitative variables such as the total number of intrusions and false-positive errors are only moderately correlated between the two tests. Comparing a revised form of the HVLT (HVLT-R; Benedict, Schretlen, Groninger, & Brandt, 1998) with the CVLT word list, Lacritz, Cullum, Weiner, & Rosenberg (2001) also found that qualitative aspects of learning (such as perseverative responses) were less correlated with one another than were learning, recall, or recognition variables. Both studies showed that the CVLT induces more errors than other verbal learning tests. Thus, further studies are needed to compare qualitative variables from word lists with each other, as well as to adapt the transformation formulae found in this study to other (amnestic) patient groups and to healthy individuals.

Funding

This work was supported by the Swiss National Science Foundation [NF 32-49107.96], a Swiss National Science Foundation Ambizione Fellowship (K.I.T.), and the Novartis Foundation, Basel, Switzerland.

Conflict of Interest

None declared.

References

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association Press.

Bäckman, L., Small, B. J., & Fratiglioni, L. (2001). Stability of the preclinical episodic memory deficit in Alzheimer's disease. *Brain*, *124 (Pt 1)*, 96–102.

Benedict, R. H. B., Schretlen, D., Groninger, L., & Brandt, J. (1998). Hopkins Verbal Learning Test-Revised: normative data and analysis of inter-form and test–retest reliability. *The Clinical Neuropsychologist*, *12*, 43–55.

Berres, M., Monsch, A. U., Bernasconi, F., Thalmann, B., & Stähelin, H. B. (2000). Normal ranges of neuropsychological tests for the diagnosis of Alzheimer's disease. *Studies in Health Technology and Informatics*, *77*, 195–199.

Berres, M., Volland, R., Beck, I., & Monsch, A. U. (2011). Orthogonal regression and related methods for bivariate data sets. Zürich: CEN.

Berres, M., Zehnder, A., Bläsi, S., & Monsch, A. U. (2008). Evaluation of diagnostic scores with adjustment for covariates. *Statistics in Medicine*, *27 (10)*, 1777–1790.

Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *The Clinical Neuropsychologist*, *5*, 125–142.

Butters, N., Delis, D. C., & Lucas, J. A. (1995). Clinical assessment of memory disorders in amnesia and dementia. *Annual Review of Psychology*, *46*, 493–523.

Caselli, R. J., Reiman, E. M., Osborne, D., Hentz, J. G., Baxter, L C., Hernandez, J. L., et al. (2004). Longitudinal changes in cognition and behavior in asymptomatic carriers of the APOE e4 allele. *Neurology*, *62*, 1990–1995.

Chen, P., Ratcliff, G., Belle, S. H., Cauley, J. A., DeKosky, S.T., & Ganguli, M. (2001). Patterns of cognitive decline in presymptomatic Alzheimer disease: A prospective community study. *Archives of General Psychiatry*, *58*, 853–858.

Collie, A., & Maruff, P. (2000). The neuropsychology of preclinical Alzheimer's disease and mild cognitive impairment. *Neuroscience and Biobehavioral Reviews*, *24*, 365–374.

Cullum, C. M., Filley, C. M., & Kozora, E. (1995). Episodic memory function in advanced aging and early Alzheimer's disease. *Journal of the International Neuropsychological Society*, *1*, 100–103.

Deary, I. J., Whiteman, M. C., Pattie, A., Starr, J. M., Hayward, C., Wright, A. F., et al. (2004). Apolipoprotein E gene variability and cognitive functions at age 79: A follow-up of the Scottish Mental Survey of 1932. *Psychology and Aging*, *19*, 367–371.

De Leon, M. J., George, A. E., Golomb, J., Tarshish, C., Convit, A., Kluger, A., et al. (1997). Frequency of hippocampal formation atrophy in normal aging and Alzheimer's disease. *Neurobiology of Aging*, *18 (1)*, 1–11.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Learning Test: Adult version*. San Antonio, TX: The Psychological Corporation.

Du, A. T., Schuff, N., Amend, D., Laakso, M. P., Hsu, Y. Y., Jagust, W. J., et al. (2001). Magnetic resonance imaging of the entorhinal cortex and hippocampus in mild cognitive impairment and Alzheimer's disease. *Journal of Neurology, Neurosurgery and Psychiatry*, *71*, 441–447.

Elias, F. E., Beiser, A., Wolf, P. A., Au, R., White, R., & D'Agostino, R. B. (2000). The preclinical phase of Alzheimer disease: A 22-year prospective study of the Framingham cohort. *Archives of Neurology*, *57*, 808–813.

Fillenbaum, G. G., Landerman, L. R., Blazer, D. G., Saunders, A. M., Harris, T. B., & Launer, L. J. (2001). The relationship of APOE genotype to cognitive functioning in older African-American and Caucasian community residents. *Journal of American Geriatric Society*, *49*, 1148–1155.

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.

Fox, L. S., Olin, J. T., Erblich, J., Ippen, C. G., & Schneider, L. S. (1998). Severity of cognitive impairment in Alzheimer's disease affects list learning using the California Verbal Learning Test (CVLT). *International Journal of Geriatric Psychiatry*, *13 (8)*, 544–549.

Fox, N. C., Warrington, E. K., Freeborough, P. A., Hartikainen, P., Kennedy, A. M., Stevens, J. M., et al. (1996). Presymptomatic hippocampal atrophy in Alzheimer's disease. *Brain*, *119*, 2001–2007.

Greenaway, M. C., Lacritz, L. H., Binegar, D., Weiner, M. F., Lipton, A., & Cullum, C. M. (2006). Patterns of verbal memory performance in mild cognitive impairment, Alzheimer Disease, and normal aging. *Cognitive Behavioral Neurology*, *19 (2)*, 79–84.

Grober, E., Lipton, R. B., Hall, C., & Crystal, H. (2000). Memory impairment on free and cued selective reminding predicts dementia. *Neurology*, *54 (4)*, 827–832.

Jonker, C., Launer, L. J., Hooijer, C., & Lindeboom, J. (1996). Memory complaints and memory impairment in older individuals. *Journal of the American Geriatrics Society*, *44 (1)*, 44–49.

Kaltreider, L. B., Cicerello, A. R., Lacritz, L. H., Weiner, M. F., Honig, L. S., Rosenberg, R. N., et al. (2000). Comparison of the CERAD and CVLT list-learning tasks in Alzheimer's disease. *The Clinical Neuropychologist*, *14 (3)*, 269–274.

Krasuski, J. S., Alexander, G. E., Horwitz, B., Daly, E. M., Murphy, D. G. M., Rapoport, St. I., et al. (1998). Volumes of medial temporal lobe structures in patients with Alzheimer's disease and mild cognitive impairment (and healthy controls). *Biological Psychiatry*, *43*, 60–68.

Lacritz, L. H., & Cullum, C. M. (1998). The Hopkins Verbal Learning Test and CVLT: A preliminary comparison. *Archives of Clinical Neuropsychology*, *13 (7)*, 623–628.

Lacritz, L. H., Cullum, C. M., Weiner, M. F., & Rosenberg, R. N. (2001). Comparison of the Hopkins Verbal Learning Test-Revised to the California Verbal Learning Test in Alzheimer's disease. *Applied Neuropsychology*, *8 (3)*, 180–184.

Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate analysis*. London: Academic Press.

Martin, A., & Fedio, P. (1983). Word production and comprehension in Alzheimer's disease: The breakdown of semantic knowledge. *Brain and Language*, *19*, 124–141.

Masur, D. M., Sliwinski, M., Lipton, R. B., Blau, A. D., & Crystal, H. A. (1994). Neuropsychological prediction of dementia and the absence of dementia in healthy elderly persons. *Neurology*, *44*, 1427–1432.

McDowell, B., Bayless, J. D., Moser, D. J., Meyers, J. E., & Paulsen, J. S. (2004). Concordance between the CVLT and the WMS-III word lists test. *Archives of Clinical Neuropsychology*, *19*, 319–324.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of department of health and human services task force on Alzheimer's disease. *Neurology*, *34*, 939–944.

Monsch, A. U., Bondi, M. W., Butters, N., Salmon, D. P., Katzman, R., & Thal, L. J. (1992). Comparisons of verbal fluency tasks in the detection of dementia of the Alzheimer type. *Archives of Neurology*, *49 (12)*, 1253–1258.

Monsch, A. U., Thalmann, B., Schneitter, M., Bernasconi, F., Aebi, C., Camachova, Z., et al. (2000). The Basel Study on the Elderly's search for preclinical cognitive markers of Alzheimer's disease. *Neurobiology of Aging*, *21*, 31.

Morris, J. C., Mohs, R. C., Rogers, H., Fillenbaum, G., & Heyman, A. (1988). Consortium to establish a registry for Alzheimer's disease (CERAD) clinical and neuropsychological assessment of Alzheimer's disease. *Psychopharmacological Bulletin*, *24 (4)*, 641–652.

Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S., et al. (1998). Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology*, *51 (6)*, 1546–1554.

Nebes, R. D. (1989). Semantic memory in Alzheimer's disease. *Psychological Bulletin*, *106 (3)*, 377–394.

Rempel-Clower, N. L., Zola, S. M., Squire, L. R., & Amaral, D. G. (1996). Three cases of enduring memory impairment after bilateral damage limited to the hippocampal formation. *The Journal of Neuroscience*, *16*, 5233–5255.

Rey, A. (1964). *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.

Romàn, G. C., Erkinjuntti, T., Wallin, A., Pantoni, L., & Chui, H. C. (2002). Subcortical ischaemic vascular dementia. *Lancet Neurology*, *1*, 426–436.

Romàn, G. C., Tatemichi, T. K., Erkinjuntti, T., Cummings, J. L., Masdeu, J. C., Garcia, J. H., et al. (1993). Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology*, *43 (2)*, 250–260.

Salmon, D. P., Heindel, W. C., & Lange, K. (1999). Differential decline in word generation from phonemic and semantic categories during the course of Alzheimer's disease: Implications for the integrity of semantic memory. *Journal of the International Neuropsychological Society*, *5*, 692–703.

Saxton, J., Lopez, O. L., Ratcliff, G., Culberg, C., Fried, L. P., Carlson, M. C., et al. (2004). Preclinical Alzheimer disease: Neuropsychological test performance 1.5 to 8 years prior to onset. *Neurology*, *63*, 2341–2347.

Small, B. J., Rosnick, C. B., Fratiglioni, L., & Bäckman, L. (2004). Apolipoprotein E and cognitive performance: A meta-analysis. *Psychology and Aging*, *19 (4)*, 592–600.

Squire, L. R., & Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science*, *253*, 1380–1386.

Stallings, G., Boake, C., & Sherer, M. (1995). Comparison of the California Verbal Learning Test and the Rey Auditory Verbal Learning Test in head-injured patients. *Journal of Clinical and Experimental Neuropsychology*, *17 (5)*, 706–712.

Twamley, E. W., Legendre Ropacki, S. A., & Bondi, M. W. (2006). Neuropsychological and neuroimaging changes in preclinical Alzheimer's disease. *Journal of the International Neuropsychological Society*, *12 (5)*, 707–735.

Victor, M., & Agamanolis, D. (1990). Amnesia due to lesions confined to the hippocampus: A clinical-pathologic study. *Journal of Cognitive Neuroscience*, *2*, 246–257.

Volland, R. (2011). Lineare Modelle für bivariate Daten. Master Thesis, RheinAhrCampus Remagen.

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

Weingartner, H., Kaye, W., Smallberg, S. A., Ebert, M. H., Gillin, J. C., & Sitaram, N. (1981). Memory failures in progressive idiopathic dementia. *Journal of Abnormal Psychology*, *90*, 187–196.

Welsh, K., Butters, N., Hughes, J., Mohs, R., & Heyman, A. (1991). Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. *Archives of Neurology*, *48 (3)*, 278–281.

Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L.-O., et al. (2004). Mild cognitive impairment – beyond controversies, towards a consensus: Report of the International Working Group on Mild Cognitive Impairment. *Journal of Internal Medicine*, *256*, 240–246.

Zehnder, A. E., Bläsi, S., Berres, M., Monsch, A. U., Stähelin, H. B., & Spiegel, R. (2009). Impact of APOE status on cognitive maintenance in healthy elderly persons. *International Journal of Geriatric Psychiatry*, *24*, 132–141.