Visual associations to retrieve episodic memory across healthy elderly, mild cognitive impairment, and patients with Alzheimer's disease

Meyer, S.R.A.; De Jonghe , J.F.M.; Schmand, B.; Ponds, R.W.H.M.

DOI
10.1080/13825585.2018.1475002

Publication date
2019

Document Version
Final published version

Published in
Aging, Neuropsychology, and Cognition

AGING, NEUROPSYCHOLOGY, AND COGNITION
2019, VOL. 26, NO. 3, 447–462
https://doi.org/10.1080/13825585.2018.1475002

Routledge
Taylor & Francis Group

Visual associations to retrieve episodic memory across healthy elderly, mild cognitive impairment, and patients with Alzheimer's disease

Sascha R. A. Meyer[a], Jos F. M. De Jonghe[b], Ben Schmand[c,d] and Rudolf W. H. M. Ponds[d,e]

[a]Department of Medical Psychology, School of Mental Health and Neuroscience (MHeNS), Maastricht University Medical Center, Maastricht, the Netherlands; [b]Department of Geriatric Medicine, Northwest Medical Center Location Alkmaar, Alkmaar, the Netherlands; [c]Department of Psychology, University of Amsterdam, Amsterdam, the Netherlands; [d]Department of Medical Psychology, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands; [e]Department of Medical Psychology, School for Mental Health and Neuroscience (MHeNS), Maastricht University Medical Center, Maastricht, the Netherlands

ABSTRACT

Episodic memory tests need to determine the degree to which patients with moderate to severe memory deficits can still benefit from retrieval support. Especially in the case of Alzheimer's disease (AD), this may support health care to be more closely aligned with patients' memory capacities. We investigated whether the different measures of episodic memory of the Visual Association Test-Extended (VAT-E) can provide a more detailed and informative assessment on memory disturbances across a broad range of cognitive decline, from normal to severe impairment as seen in AD, by examining differences in floor effects. The VAT-E consists of 24 pairs of black-and-white line drawings. In a within-group design, we compared score distributions of VAT-E subtests in healthy elderly controls, mild cognitive impairment (MCI), and AD ($n = 144$), as well as in relation to global cognitive impairment. Paired associate recall showed a floor effect in 41% of MCI patients and 62% of AD patients. Free recall showed a floor effect in 73% of MCI patients and 84% of AD patients. Multiple-choice cued recognition did not show a floor effect in either of the patient groups. We conclude that the VAT-E covers a broad range of episodic memory decline in patients. As expected, paired associate recall was of intermediate difficulty, free recall was most difficult, and multiple-choice cued recognition was least difficult for patients. These varying levels of difficulty enable a more accurate determination of the level of retrieval support that can still benefit patients across a broad range of cognitive decline.

ARTICLE HISTORY
Received 27 October 2017
Accepted 3 May 2018

KEYWORDS
Alzheimer's disease; mild cognitive impairment; dementia; learning and memory; assessment

Introduction

Cognitive impairment is measured with episodic memory tests, but it is unclear whether such tests are clinically meaningful in other ways, particularly in determining if patients with Alzheimer's disease (AD) can still benefit from retrieval support. Ideally, this should

CONTACT Sascha R. A. Meyer smeyer70@gmail.com

be determined across a broad range of cognitive decline, so that health care can be more closely aligned with a patient's memory capacities by providing both health-care professionals and informal caregivers with memory tools that improve the patient's daily memory functioning.

In clinical practice, most commonly used episodic memory tests are based on free recall paradigms, which are primarily used to distinguish between normal and abnormal, dementia-related memory functioning, ignoring more subtle differences in the level of intact memory processing across the broad range of global cognitive impairment in dementia patients. A more informative approach would be to use tasks with varying levels of retrieval support to assess the strength of memory processing in a more detailed way. Earlier research has shown that test paradigms that provide retrieval support by means of cued recall might improve retrieval of test material in patients with mild cognitive impairment (MCI) or AD (e.g., Greenaway et al., 2006; Tounsi et al., 1999). The retrieval of test material in these patients might further be improved if even more retrieval support is provided by using recognition paradigms (Bäckman, Jones, Berger, Laukka, & Small, 2005; Lowndes et al., 2008). In addition, research showed that free recall can be too difficult for patients with MCI or AD, thus resulting in floor effects on tests, whereas cued recall subsequently improved retrieval of episodic memory in these patients (Diesfeldt, Prins, & Lauret, 2017; Meyer et al., 2016). Therefore, several test paradigms with varying difficulty levels should be employed to cover a broad range of memory decline in these patients, and in order to enable comparison between test paradigms, these paradigms should be combined in one single test.

A well-known example of such a test is the California Verbal Learning Test (Delis, Kramer, Kaplan, & Ober, 1987), which combines free recall, cued recall and recognition paradigms. Research has shown that these measures can differentiate within a group of AD patients based on severity of cognitive impairment (Fox, Olin, Erblich, Ippen, & Schneider, 1998), thus covering a broad range of memory decline.

Since the California Verbal Learning Test is based on verbal stimuli, a similar test based on visual stimuli might be clinically useful to enable language-independent assessment of patient groups, such as dyslexic or non-native speakers. In addition, tests based on visual stimuli are particularly suitable for covering a broad range of memory decline because pictures are easier to remember than words (Shepard, 1967) and because this so-called picture superiority effect is applicable to normal controls for free recall (Paivio, Rogers, & Smythe, 1968), paired associate recognition (Hockley, 2008), and recognition (Shepard, 1967), as well as to patients with MCI or AD for recognition (Alley, Gold, & Budson, 2009). However, current visual memory tests often use abstract designs as learning material, e.g., Wechsler Memory Scale-IV subtest Visual Reproduction (PsychCorp, 2009; Wechsler, 1945), Rey's Complex Figure Test (Osterreith, 1944; Rey, 1941), and Benton Visual Retention Test (Sivan, 1992). Because these abstract designs have to be recalled by drawing the design from memory, this may complicate the interpretation since failure may be due to visuoconstructive disability. Also, since functioning in daily life mostly does not concern abstract designs, such tests seem less applicable to patients' daily memory functioning. Moreover, because the memory deficit in AD has mostly been studied by means of word-list learning (Lezak, Howieson, Bigler, & Tranel, 2012), a test that is based on pictures related to daily life may be a valuable addition for studying memory in MCI and AD.

The Visual Association Test (VAT; Lindeboom & Schmand, 2003; Lindeboom, Schmand, Meyer, & De Jonghe, 2014) is an episodic memory test based on pictures of

everyday objects, animals, plants, or food. Various studies showed that the VAT has a high validity in predicting progression to AD (Dierckx et al., 2009; Fuchs, Wiese, Altiner, Wollnyu, & Penzek, 2011; Lindeboom, Schmand, Tulner, Walstra, & Jonker, 2002; Reesink et al., 2010; Van Der Vlies et al., 2009). We previously extended the VAT by doubling the number of test items and adding a free recall condition (Meyer, De Jonghe, Schmand, & Ponds, 2017), thus constructing the measures of memory of the Visual Association Test-Extended (VAT-E; Meyer & De Jonghe, 2017), i.e., paired associate recall, free recall and multiple-choice cued recognition, which, in conjunction with each other, may be suitable for covering a broad range of memory decline.

The VAT-E is based on visual associations shown during an incidental learning task (Craik & Lockhart, 1972; Tulving & Madigan, 1970). By means of the visual association, a picture that the patient is asked to remember is encoded together with a specific cue, which subsequently can be presented to maximize retrieval. By varying in the level of retrieval support, also the level of difficulty of the measures of memory varies, which in turn enables a broad range of memory decline to be covered. The VAT-E consists of pairs of pictures that are associated through an interaction, e.g., a dog is depicted riding a bicycle (see Figure 1). The first measure of memory, paired associate recall, is thought to be of intermediate difficulty, as it presents a cue, e.g., a picture of a bicycle, to facilitate recall (see Figure 1). The second measure of memory, free recall, is thought to be the most difficult, as no retrieval support is presented. The third measure of memory, multiple-choice cued recognition, is thought to be the least difficult, as it not only presents a cue, e.g., a picture of a bicycle, but also facilitates recognition by presenting the target and three alternatives (see Figure 1).

The purpose of our study was to investigate if the VAT-E covers a broad range of memory decline by examining the difficulty levels of the measures of episodic memory in patients with MCI or AD across the range of severity of global cognitive impairment. These varying levels of difficulty would enable a more accurate determination of the level of retrieval support that can still benefit patients across a broad range of cognitive decline. The difficulty level of a memory test can be evaluated by examining floor effects, with larger floor effects indicating increasingly difficult measures (Cohen, Swerdlik, & Sturman, 2013). We expected the more difficult measures to show larger floor effects in patient groups with more severe global cognitive impairment. We hypothesized that the VAT-E measures of memory cover a broad range of cognitive decline in patients with MCI or AD, since they do not only establish,

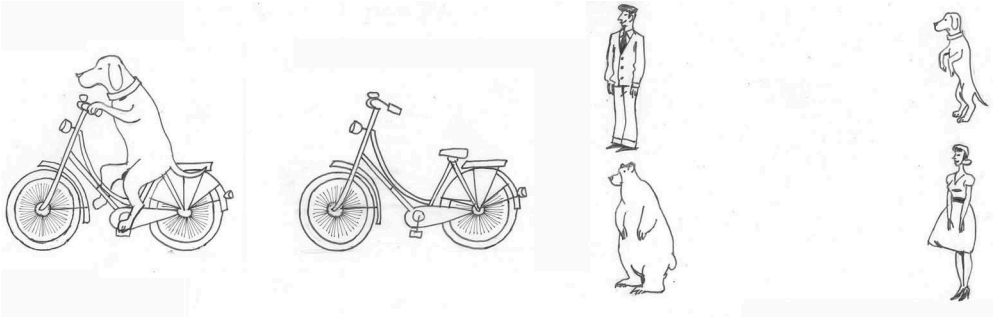

Figure 1. Sample of visual association (left), cue (middle), and multiple-choice answer (right).

by means of free recall, whether a patient has an episodic memory impairment, but also give information, by means of paired associate recall and multiple-choice cued recognition, about the degree to which a patient can still benefit from retrieval support. We expected that paired associate recall would show an intermediate floor effect, as it is of intermediate difficulty for patient groups, while free recall was expected to show a large floor effect, as it is the most difficult task for patient groups, and multiple-choice cued recognition was expected to show no floor effect, as it is the least difficult task for patient groups.

Methods

Participants

Our study sample consisted of 44 healthy elderly controls, 74 MCI patients, and 26 mild AD patients. The MCI group consisted of 59 patients with amnestic MCI and 15 with amnestic MCI multiple domain (Petersen, 2004). Hearing and vision was intact or corrected to normal for all participants. We recruited patients from the outpatient memory clinic at the department of geriatric medicine of the Northwest Medical Center, a large teaching hospital in Alkmaar, the Netherlands. Healthy elderly controls were recruited among the informal caregivers of the participating outpatients, all of whom were close family members or spouses of the patients. Since all informal caregivers cared for the patients on a regular basis, they were extensively interviewed by the clinical neuropsychologist regarding the medical history of the patient. If the informal caregivers appeared clinically normal, had no complaints of memory, and were generally unaffected in their activities of daily living, they were asked to participate in the study as healthy elderly controls.

Participants were included if they were 65 years or older. Participants were excluded if they had severe traumatic brain damage or brain disease (such as brain tumor, epilepsy, multiple sclerosis, Parkinson's disease), a severe psychiatric disorder, delirium or a history of alcohol or drug abuse. Healthy elderly controls were included if they scored above published cutoff scores, i.e., 13 or higher, on the Cognitive Screening Test (CST; De Graaf & Deelman, 1991). Healthy elderly controls underwent a screening interview concerning the exclusion criteria and were excluded, if appropriate. In addition, they were excluded if their Rey Auditory Verbal Learning Test delayed recall score was at percentile 5 or below corrected for gender, age, and education.

MCI and AD diagnoses were based on internationally accepted criteria (*Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*, Text Revision; American Psychiatric Association, 2000; Dubois et al., 2007; Petersen et al., 2001). The patients' diagnostic program included a comprehensive geriatric assessment, laboratory results, magnetic resonance imaging or computed tomography scanning, and a standardized neuropsychological examination. The final MCI or AD diagnosis was made after discussion of all clinical data by a multidisciplinary team consisting of five geriatricians, a neurologist, a clinical neuropsychologist and a geriatric psychiatrist. The study was approved by the local medical ethical committee. Written informed consent was obtained from all participants.

Design and procedure

In a within-group design, we compared VAT-E scores for paired associate recall, free recall, and multiple-choice cued recognition in groups of healthy elderly controls, MCI patients, and AD patients, in relation to the participants' severity of global cognitive impairment as measured with the CST. As an additional baseline measure, all participants took the Rey Auditory Verbal Learning Test (Rey, 1964; Schmidt, 1996) delayed recall trial. For patients, administration of the VAT-E, the CST, and the Rey Auditory Verbal Learning Test was part of the standardized neuropsychological examination of the hospital, which besides these tests, consisted of several tests and questionnaires, i.e., Cambridge Cognitive Examination subtest perception (Roth, Huppert, Mountjoy, & Tym, 1998); Animal Gnosis-card, Amsterdam Dementia-Screening Test subtests orientation, copy, and meander (Lindeboom & Jonker, 1989); Trail Making Test (Reitan & Wolfson, 1985); Stroop Color–Word Test (Hammes, 1971); Clock drawing, Groninger Intelligence Test 2 subtests fluency animals, fluency professions, word meaning, and additional subtests if appropriate (Luteijn & Barelds, 2004); Rey's Complex Figure Test subtests copy and delayed recall, Geriatric Depression Scale (Sheik & Yesavage, 1986); Informant Questionnaire on Cognitive Decline in the Elderly (Jorm & Jacomb, 1989); Neuropsychiatric Inventory (Cummings et al., 1994); Lawton Instrumental Activities of Daily Living (Lawton & Brody, 1969); KATZ activities of daily living (Katz, Ford, Moskowitz, Jackson, & Jaffe, 1963); Clinical Dementia Rating (Hughes, Berg, Danziger, Coben, & Martin, 1982; Morris, 1993); and subtests of the Behavioural Assessment of the Dysexecutive Syndrome if appropriate (Wilson, Alderman, Burgess, Emslie, & Evans, 1996). Since healthy elderly controls were the informal caregivers of the patients, they did not undergo the standardized neuropsychological examination but only the VAT-E, the CST, and the Rey Auditory Verbal Learning Test. To avoid interference between these tests, they were administered in a fixed order, i.e., CST, VAT-E, and Rey Auditory Verbal Learning Test, and during the time interval between the immediate and delayed conditions of the VAT-E and the Rey Auditory Verbal Learning Test, participants were requested to complete questionnaires. However, if, in spite of this, there would have been interference, this would be equal for all participants due to the fixed order. Healthy elderly controls completed these tests at the same time as the patients, but in a separate room.

Since the VAT-E was not used to determine the patients' diagnosis, we removed the VAT-E score sheet from the patient's hospital dossier to ensure that the clinical neuropsychologist was blinded to the VAT-E score. The CST could not be removed from the patient's hospital dossier, as it is part of the hospital's standardized neuropsychological examination.

Measures

The VAT-E (Meyer & De Jonghe, 2017) is based on items of the Visual Association Test (VAT; Lindeboom & Schmand, 2003; Lindeboom et al., 2014) and consists of the joined VAT 12-item versions AB and CD. Only the CD version has a four-alternative forced-choice cued recognition trial. This version was used for the multiple-choice cued recognition trial of the VAT-E.

The VAT-E measures of episodic memory, paired associate recall, free recall, and multiple-choice cued recognition are based on 24 pictures of black-and-white line

drawings. The stimuli are learned in two learning trials and two recognition trials, i.e., immediate recognition and delayed recognition. First, each learning trial presents the same 24 pictures one by one, e.g., a bicycle, which the subject is instructed to look at and remember. In the immediate recognition trial, each picture from the learning trials is linked with an associated picture through an interaction, e.g., a dog riding a bicycle (see Figure 1). The subject is instructed to name both pictures, thus inducing incidental learning of the association, and to point out the picture from the learning trials, e.g., the bicycle. After 15 min, delayed recognition is administered. During this time interval, subjects can complete nonvisual tests, non-memory tests, or questionnaires. Delayed recognition is administered in the same way as immediate recognition, i.e., the subject is instructed to name both pictures and point out the picture from the learning trials. In this way, the subject is being tested twice on recall of the learning trial pictures. The immediate and delayed recognition trials are employed as performance validity measures (Meyer et al., 2017).

Next, the paired associate recall trial is administered, in which the pictures from the learning trials are shown again, but now one by one as cues, e.g., the bicycle (see Figure 1), and the subject is instructed to recall the associated picture, e.g., the dog. The paired associate recall scores range from 0 to 24. Next, the free recall trial is administered, in which the subject is instructed to recall all cues and all associated pictures without being shown any of the pictures. The free recall scores range from 0 to 48. Finally, the multiple-choice cued recognition trial is administered, in which the last 12 pictures from the learning trials are shown again one by one as cues, and the subject is instructed to recognize and point out the associated picture out of four alternatives (see Figure 1). The multiple-choice cued recognition scores range from 0 to 12.

During all trials, the test administrator gives feedback regarding correct or incorrect responses on each item. If, during paired associate recall, the subject gives an incorrect answer, the test administrator gives the correct answer so that the exposure to the cue and the associated picture is equal for all subjects before the free recall trial. An incorrect response is also corrected during multiple-choice cued recognition, but not during free recall. The VAT-E measures of memory are highly predictive of MCI and AD, with areas under the curve from .81 to .99 for MCI, and from .97 to .99 for AD (Meyer & De Jonghe, 2017). Normative data for the VAT-E measures of memory were established (Meyer & De Jonghe, 2017).

The severity of global cognitive impairment was measured with the CST (De Graaf & Deelman, 1991), which consists of 20 questions relating to orientation in time, place, and person, and elementary factual knowledge, e.g., "when was the second World War?". The scores range from 0 to 20. High scores indicate lack of impairment and lower scores indicate more severe impairment. The CST is a Dutch version of the short portable mental status questionnaire (Pfeiffer, 1975) and the mental status questionnaire (Kahn, Goldfarb, Pollack, & Peck, 1960). The CST correlates 0.86 with the Mini-Mental State Examination (Dautzenberg, Schmand, Vriens, Deelman, & Hooijer, 1991; Folstein, Folstein, & McHugh, 1975). In a study based on a large sample, n = 4051, the CST was compared to the Mini-Mental State Examination and was found to be equally predictive of dementia, with areas under the curve of .91 (95% confidence interval [CI] = .88–.93) and .92 (95% CI = .90–.94) respectively (Schmand, Deelman, Hooijer, Jonker, & Lindeboom, 1996). In another study based on a large sample, n = 761, the CST was also highly predictive of dementia, with an area under the

curve of .94 (Van Toutert, Diesfeldt, & Hoek, 2016). The reliability of the CST is high, with Cronbach's alpha being 0.89 (Dautzenberg et al., 1991) and test–retest reliability being 0.80 (Schmand et al., 1996).

This study used the Dutch version of the Rey Auditory Verbal Learning Test (Saan & Deelman, 1986), which presents a list of 15 unrelated words in five consecutive learning trials. After each presentation and after a 20-min time interval, i.e., delayed recall, these words have to be recalled. The delayed recall scores range from 0 to 15.

The level of education was scored with a Dutch system (Verhage, 1964) that is similar to the International Standard Classification of Education 1997 scales (UNESCO, 2006). The scores range from 1 to 7, i.e., preprimary education, primary education, lower secondary education, upper secondary education, postsecondary non-tertiary education, first-stage tertiary education, and second-stage tertiary education respectively.

Data analysis

The effects of demographic characteristics on VAT-E paired associate recall, free recall, and multiple-choice cued recognition were analyzed by means of Pearson's chi-square test, Kruskal–Wallis tests, Mann–Whitney tests, *t*-tests and Spearman's and Pearson's correlations, as appropriate. Also, we analyzed differences between groups on VAT-E paired associate recall, free recall, and multiple-choice cued recognition, as appropriate. VAT-E scores of healthy elderly controls, MCI patients, and AD patients were pooled. VAT-E scores were plotted against the range of global cognitive impairment, i.e., the participants' CST scores. Hierarchical multiple regression analysis was used to determine whether a linear or quadratic model fitted the data best. To predict VAT-E test performances, we entered the CST score, representing a linear model, at step 1, and entered the squared CST score, representing a quadratic model, at step 2. This resulted either in a linear model as best fit for the data if the total variance explained was significantly increased by the CST score alone or a quadratic model as best fit for the data if the total variance explained was significantly increased by the CST score together with the squared CST score.

If the best fit for the data is a quadratic model, this means that VAT-E scores level off as severity of global cognitive impairment increases, thus indicating a floor effect. The range of such a floor effect was defined by scores below the VAT-E score that was obtained from entering the minimum CST score into the quadratic equation and calculating the outcome. Subsequently, we calculated the percentage of subjects scoring within the range of the floor effect. Preliminary analyses showed that the assumption of multicollinearity was violated, but this was as expected, because the two predictors in the model are based on the same variable, i.e., the CST score. Also, the assumption of linearity was violated, as expected, for the quadratic model.

Results

The demographic characteristics and test performances per group are shown in Table 1. The groups differed in gender ($\chi^2(2) = 15.99$, $p < .001$) and age ($H(2) = 18.74$, $p < .001$). These differences had no effect on VAT-E scores. For paired associate recall, healthy elderly controls ($Mdn = 20.0$) scored higher than MCI patients ($Mdn = 6.0$) ($U = 398.5$, $p < .001$), and MCI patients

Table 1. Demographic characteristics and test performances per group.

	Healthy elderly controls (n = 44)	MCI (n = 74)	AD (n = 26)
Female (%)	84.1	47.3	53.8
Age	73.0 (5.6)	76.5 (5.6)	79.5 (6.7)
Education	4.6 (1.1)	4.2 (1.1)	4.3 (1.0)
VAT-E paired associate recall	18.1 (4.8; 23.3; 5–24)	7.9 (6.5; 42.6; 0–23)	4.3 (4.3; 18.7; 0–18)
VAT-E free recall	22.8 (8.0; 64.1; 9–48)	6.0 (4.0; 15.6; 0–17)	3.6 (4.2; 17.8; 0–16)
VAT-E multiple-choice cued recognition	11.9 (0.3; 0.1; 11–12)	9.6 (2.7; 7.3; 2–12)	7.2 (3.5; 11.9; 1–12)
CST	19.2 (1.2; 1.3; 15–20)	16.0 (2.0; 4.2; 8–20)	13.6 (2.7; 7.5; 6.5–18)
Rey Auditory Verbal Learning Test delayed recall	8.5 (2.5; 6.3; 5–13)	2.0 (1.8; 3.3; 0–6)	0.9 (1.5; 2.3; 0–6)

MCI: mild cognitive impairment; AD: Alzheimer's disease; n: number of subjects; female (proportion of females); age in years (mean and standard deviation); education score based on the Verhage system (mean and standard deviation); VAT-E: Visual Association Test-Extended (mean, standard deviation, variance, and range of scores); CST: Cognitive Screening Test (mean, standard deviation, variance and range of scores); Rey Auditory Verbal Learning Test (mean, standard deviation, variance and range of scores).

scored higher than AD patients (Mdn = 3.5) (U = 636.0, p < .05). For free recall, healthy elderly controls (Mdn = 21.0) scored higher than MCI patients (Mdn = 5.0) (U = 52.0, p < .001), and MCI patients scored higher than AD patients (Mdn = 3.0) (U = 539.5, p < .01). For multiple-choice cued recognition, healthy elderly controls (Mdn = 12.0) scored higher than MCI patients (Mdn = 10.5) (U = 668.0, p < .001), and MCI patients scored higher than AD patients (Mdn = 7.5) (U = 543.0, p < .01).

Table 2 and Figures 2–4 show the test performances of VAT-E paired associate recall, free recall, and multiple-choice cued recognition in relation to severity of global cognitive impairment. For paired associate recall, the explained variance in test performance was increased by the squared CST scores. Thus, the best fit for the paired associate recall data was a quadratic regression model (F (2, 141) = 46.87, p < .001) (see Figure 2), which indicated a floor effect. The range of this floor effect comprised scores of 4 or below, which were scored by 41% of MCI patients and 62% of AD patients. No healthy elderly controls scored within the range of this floor effect. For free recall, the explained variance in test performance was increased by the squared CST scores. Thus, the best fit for the free recall data was a quadratic regression model (F (2, 140) = 68.22, p < .001) (see Figure 3), which indicated a floor effect. The range of this floor effect comprised scores of 7 or below, which were scored by 73% of MCI patients and 84% of AD patients. No healthy elderly controls scored within the range of this floor effect. For multiple-choice cued recognition, the explained variance in test performance was not increased by the squared CST scores. Thus, the best fit for the multiple-choice cued recognition data was a linear regression model (F (1, 142) = 48.32, p < .001) (see Figure 4), which indicated that multiple-choice cued recognition did not show a floor effect.

Discussion

This study aimed to investigate if the VAT-E covers a broad range of memory decline by examining the difficulty levels of several episodic memory measures across a wide range of global cognitive impairment. We compared the outcomes of VAT-E subtests in normal controls, patients with amnestic MCI, and patients with mild AD. Free recall was associated

Table 2. Summary of hierarchical regression analysis for CST variables predicting VAT-E paired associate recall, free recall, and multiple-choice cued recognition total scores (n = 144).

VAT-E	Variable	*B*	*SE B*	β
Paired associate recall	Step 1 (linear model)			
	Constant	−17.16	3.2	
	CST	1.66	0.19	.60***
	Step 2 (quadratic model)			
	Constant	17.69	11.03	
	CST	−3.03	1.44	−1.08*
	CST squared	0.15	0.05	1.69**
Free recall	Step 1 (linear model)			
	Constant	−26.40	3.85	
	CST	2.24	0.23	.64***
	Step 2 (quadratic model)			
	Constant	34.50	12.79	
	CST	−5.97	1.67	−1.69***
	CST squared	0.27	0.05	2.35***
Multiple-choice cued recognition	Step 1 (linear model)			
	Constant	1.17	1.27	
	CST	0.53	0.08	.50***
	Step 2 (quadratic model)			
	Constant	1.65	4.59	
	CST	0.46	0.60	.44
	CST squared	0.00	0.02	.06

PA: R^2 = .353 for Step 1; ΔR^2 = .046 for Step 2 (F change (1, 141) = 10.81, p < .01); FR: R^2 = .405 for Step 1; ΔR^2 = .089 for Step 2 (F change (1, 140) = 24.58, p < .001); MC: R^2 = .254 for Step 1; ΔR^2 = .0001 for Step 2 (F change (1, 141) = 0.01, p = .912).
*p < .05; **p < .01; ***p < .001.
CST: Cognitive Screening Test; VAT-E: Visual Association Test-Extended; n: number of subjects.

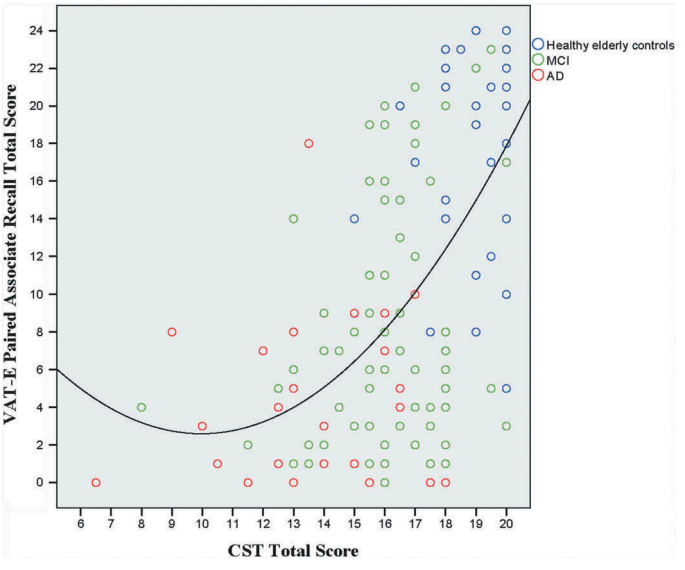

Figure 2. VAT-E paired associate recall scores compared to CST scores (n = 144). VAT-E: Visual Association Test-Extended; CST: Cognitive Screening Test; *n:* number of subjects; MCI: mild cognitive impairment; AD: Alzheimer's disease.

with floor effects in both patient groups, but paired associate recall showed a smaller floor effect, and it was absent in the multiple-choice cued recognition trial. Healthy controls scored well above the floor levels. These results indicate that if strong links are provided

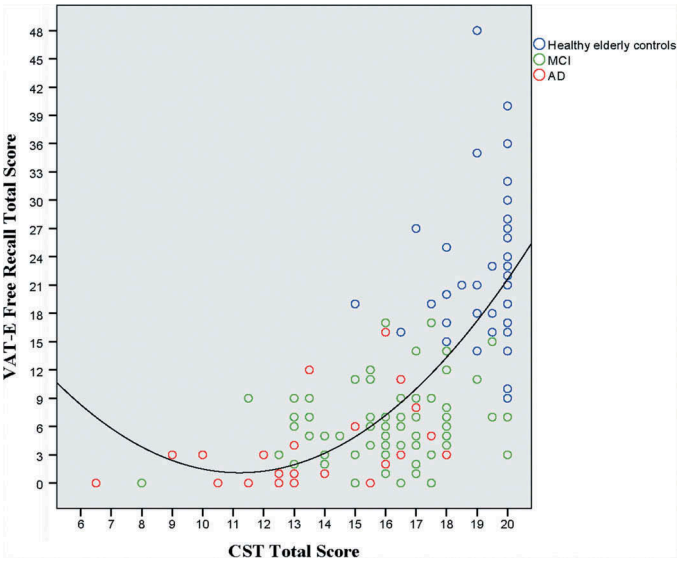

Figure 3. VAT-E free recall scores compared to CST scores (n = 144). VAT-E: Visual Association Test-Extended; CST: Cognitive Screening Test; *n:* number of subjects; MCI: mild cognitive impairment; AD: Alzheimer's disease.

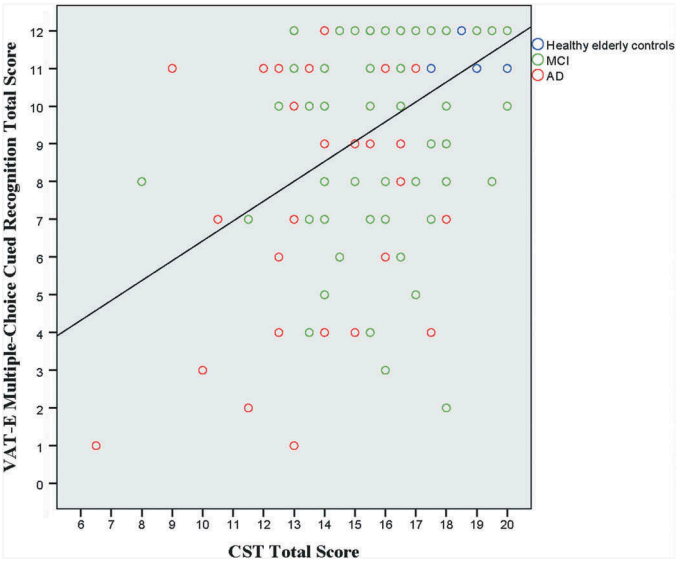

Figure 4. VAT-E multiple choice cued recognition scores compared to CST scores (n = 144). VAT-E: Visual Association Test-Extended; CST: Cognitive Screening Test; *n*: number of subjects; MCI: mild cognitive impairment; AD: Alzheimer's disease.

between cues and to be remembered material, as is the case with the VAT-E visual associations paradigm, patients with AD remember more as more retrieval support is provided. As is evident by our findings that while free recall of the visual associations was most difficult for patients with MCI or mild AD, cued recall of these visual associations was of

intermediate difficulty, and recognizing newly learned material based on cues was easiest for these patients. Thus, in conjunction with each other, the VAT-E measures can provide insight into the specific degree to which patients within a broad range of cognitive decline can still benefit from retrieval support.

Strong points of our study are that it explores a broad range of episodic memory impairment, from normal cognitive functioning via MCI to the more severe impairment of dementia, and that we used a regression-based method to examine episodic memory in relation to cognitive decline. The use of hierarchical multiple regression analysis enabled us to compare VAT-E episodic memory scores, not only in healthy elderly controls, MCI patients and AD patients, but also in relation to global cognitive impairment. Since we compared paired associate recall, free recall, and multiple-choice cued recognition using the same groups, we consider our findings to be robust.

Our results are in line with previous research. While some test paradigms, such as free recall, are well suited for measuring mild memory impairment, other paradigms, such as paired associate recall or recognition, may be preferred for measuring more severe impairment in AD (Bäckman et al., 2005; Greenaway et al., 2006; Lowndes et al., 2008; Tounsi et al., 1999). The VAT-E applies these memory paradigms by providing visual associations in which specific cues are encoded along with associated pictures during incidental learning, and by providing various degrees of retrieval support during testing (Craik & Lockhart, 1972; Tulving & Madigan, 1970). Our results indicate that within the VAT-E testing paradigm, patients with AD remember more as more retrieval support is provided, which is in line with recent findings indicating that previously forgotten information can be retrieved in early AD (Roy et al., 2016).

The visual associations enhance the engagement with the test material, and thus they support the establishment of longer lasting, and stronger memory traces (Craik & Lockhart, 1972; Tulving & Madigan, 1970). This is in line with previous research. First, in contrast to words, pictures are mostly coded visually as well as verbally. Due to this dual coding process (Paivio, 1969, 1991), the retrieval of pictures is easier than the retrieval of words (Alley et al., 2009; Hockley, 2008; Paivio et al., 1968; Shepard, 1967). Second, the interaction depicted in the visual association creates a stronger link between the cue and the associated picture than a mere side-by-side presentation of the same pictures. In other words, cued recall of the associations is improved within the incidental learning paradigm compared to cued recall of the separate pictures within an intentional learning paradigm (Lindeboom et al., 2002). Third, the depicted interactions are unexpected, and thus they are elaborated on more than expected interactions (Birngruber, Schröter, & Ulrich, 2014; Ulrich, Nitschke, & Rammsayer, 2006). In addition, providing an occasion of incidental learning gives the test administrator more control over how the subject processes the material, as opposed to intentional learning situations, in which the subject could use a mnemonic strategy that remains hidden to the test administrator (Craik & Lockhart, 1972). That is, as subjects are not explicitly instructed to learn but to perform another task, i.e., picture naming, they will not apply an unknown coding strategy. Thus, encoding will be more equal among subjects. Furthermore, the encoding is facilitated by the dual coding of pictures, the depicted interaction, and the unexpected nature of the association.

A potential limitation of this study is that global cognitive impairment was determined by means of a screening test, i.e., the CST, which might be less sensitive to MCI

than to dementia. However, sensitivity is based on cutoff scores, which we did not use in our analysis. Instead, we used the entire measurement range of the CST. Therefore, if this had had an effect on our results, then it would have been evident in a restriction of range for the MCI patients. However, as Table 1 shows, variances of CST scores increased considerably from healthy elderly controls to MCI patients to AD patients, i.e., more than 200% and 80%, respectively. Therefore, we consider the CST to be an adequate tool for determining the range of global cognitive impairment in these groups.

Another limitation is that the clinical neuropsychologist was not blinded to the CST score during the diagnostic process. This was due to the fact that the CST is part of the hospital's standardized neuropsychological examination, and thus it could not be removed from the patient's hospital dossier. However, the CST played only a very small part in the diagnosis of the patients, because it is a screening test, and the diagnostic program of the patients was far more extensive than the CST alone, as it included a comprehensive geriatric assessment, laboratory results, magnetic resonance imaging or computed tomography scanning, and a standardized neuropsychological examination consisting of several tests relating to relevant cognitive functions. Besides, the study protocol did ensure that the clinical neuropsychologist was blinded to the VAT-E scores, which were not used in the diagnostic process, and therefore we consider our results to be robust.

Another possible limitation is that we investigated patients with MCI and mild AD, and thus we did not investigate episodic memory impairment in AD for the whole range of cognitive decline. However, clinicians who want to determine the degree to which patients can still benefit from retrieval support will generally want to do so when the first symptoms of the disease come to the fore, i.e., during the early stages of the disease, which are included in our study.

To determine whether a patient with an episodic memory impairment can still benefit from retrieval support, one has to know not only the norm score but also the level of difficulty of the measures used. In addition to the norms provided in the manual of the VAT-E (Meyer & De Jonghe, 2017), our analyses show that in relation to the severity of global cognitive impairment, a large proportion of patients performed at floor level on free recall and a smaller proportion of patients performed at floor level on paired associate recall. Only on the least difficult measure, multiple-choice cued recognition, no patient performed at floor level. Thus, if a patient scores within the impaired range on free recall, this establishes an episodic memory impairment, but it gives no information about the ability of a patient to still benefit from retrieval support. For this, the clinician should also inspect the scores on paired associate recall and multiple-choice cued recognition. If, in addition to free recall, a subject benefits from the retrieval support of paired associate recall and even more so from the retrieval support of multiple-choice cued recognition, this may support health care to be more closely aligned with patients' memory capacities.

If the VAT-E indicates that a patient can still benefit from retrieval support, health-care professionals and informal caregivers can be provided with memory tools that improve memory performance in patients' daily functioning. Indeed, previous research showed that learning methods based on cues are effective in mild AD patients (Bier et al., 2008). A patient might be helped to learn to associate activities of daily living with icons, which subsequently could be recognized by that patient in a day planner. Also, health-care professionals and informal caregivers could be given tools for communicating with

patients, as they could be coached how to align their communication style with the retrieval support the patient needs, by using cues or multiple-choice options in conversation. In addition, we consider the VAT-E to be more suitable for elderly patients than the frequently used word-list learning tests, because it is based on visual associations which are encoded by means of incidental learning, making it less taxing for these patients. Also, as opposed to most other visual memory tests, the VAT-E tests memory for pictures of everyday objects, animals, plants, or food instead of abstract designs, which makes it more applicable to patients' daily memory functioning.

Episodic memory tests are needed that enable clinicians to determine the degree to which patients with MCI or AD can still benefit form retrieval support across a broad range of cognitive decline. Our findings show that the VAT-E covers a broad range of episodic memory decline in patients. In contrast to most other memory tests, the VAT-E uses visual associations and incidental learning to establish longer lasting and stronger memory traces. By subsequently varying retrieval support, we constructed measures of varying levels of difficulty. These varying levels of difficulty enable a more accurate determination of the level of retrieval support that can still benefit patients across a broad range of cognitive decline. Providing patients with the right amount of retrieval support may improve their daily functioning.

Acknowledgments

We thank the department of geriatric medicine of the Northwest Medical Center Location Alkmaar, the Netherlands, for enabling us to do our research.

Disclosure statement

S.R.A.M. and J.F.M.J. are the authors of the Visual Association Text-Extended.

Funding

This work was supported by Noordwest Academie [grant number FIO1308].

References

Alley, B. A., Gold, C. A., & Budson, A. E. (2009). The picture superiority effect in patients with Alzheimer's disease and mild cognitive impairment. *Neuropsychologia*, *47*, 595–598.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington DC: Author.

Bäckman, L., Jones, S., Berger, A., Laukka, E. J., & Small, B. J. (2005). Cognitive impairment in preclinical Alzheimer's disease: A meta-analysis. *Neuropsychology*, *19*, 520–531.

Bier, N., van der Linden, M., Gagnon, L., Desrosiers, J., Adam, S., Louveaux, S., & Saint-Mleux, J. (2008). Face-name association learning in early Alzheimer's disease: A comparison of learning methods and their underlying mechanisms. *Neuropsychological Rehabilitation*, *18*, 343–371.

Birngruber, T., Schröter, H., & Ulrich, R. (2014). Duration perception of visual and auditory oddball stimuli: Does judgement task moderate the temporal oddball effect? *Attention, Perception, & Psychophysics*, *76*, 814–828.

Cohen, R. J., Swerdlik, M. E., & Sturman, E. D. (2013). *Psychological testing and assessment. An introduction to tests and measurement* (8th ed., pp. 256, 259). New York: McGraw-Hill.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.

Cummings, J. L., Mega, M., Gray, K., Rosenberg-Thompson, S., Carusi, D. A., & Gornbein, J. (1994). The Neuropsychiatric inventory: Comprehensive assessment of psychopathology in dementia. *Neurology*, *44*, 2308–2314.

Dautzenberg, P. L. J., Schmand, B., Vriens, M. T. S., Deelman, B. G., & Hooijer, C. (1991). De validiteit van de cognitieve screening-test en de ´mini-mental state examination´ bij een groep oudere ziekenhuispatiënten [Study of the validity of the cognitive screening test and the mini-mental state examination in an older hospital population]. *Nederlands Tijdschrift Voor Geneeskunde*, *135*, 850–855.

De Graaf, A., & Deelman, B. G. (1991). *Cognitive Screening Test. Manual for administration and scoring*. Lisse: Swets en Zeitlinger bv.

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

Dierckx, E., Engelborghs, S., de Raedt, R., van Buggenhout, M., de Deyn, P. P., Verté, D., & Ponjaert-Kristoffersen, I. (2009). Verbal cued recall as a predictor of conversion to Alzheimer´s disease in mild cognitive impairment. *International Journal of Geriatric Psychiatry*, *24*, 1094–1100.

Diesfeldt, H., Prins, M., & Lauret, G., (2017). De visuele associatietest (VAT) als instrument voor de ouderenpsycholoog [The Visual Association Test to study episodic memory in clinical geriatric psychology]. *Tijdschrift Voor Gerontologie En Geriatrie*, Advanced online publication. doi:10.1007/s12439-017-0231-7

Dubois, B., Feldman, H. H., Jacova, C., DeKosky, S. T., Barberger-Gateau, P., Cummings, J., & Scheltens, P. (2007). Research criteria for the diagnosis of Alzheimer´s disease: Revising the NINCDS-ADRDA criteria. *Lancet Neurology*, *6*, 734–746.

Folstein, M. F., Folstein, S. F., & 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*, 544–549.

Fuchs, A., Wiese, B., Altiner, A., Wollnyu, A., & Penzek, M. (2011). Cued recall and other cognitive tasks to facilitate dementia recognition in primary care. *Journal of American Geriatric Society*, *60*, 130–135.

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*, 79–84.

Hammes, J. G. W. (1971). *De Stroop Kleur-Woord Test. Handleiding. [The Stroop Colour-Word Test. Manual]*. Amsterdam: Pearson Assessment and Information B.V.

Hockley, W. E. (2008). The picture superiority effect in associative recognition. *Memory & Cognition*, *36*, 1351–1359.

Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A., & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *The British Journal of Psychiatry*, *140*, 566–572.

Jorm, A. F., & Jacomb, P. A. (1989). The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): Socio-demographic correlates, reliability, validity and some norms. *Psychological Medicine*, *19*, 1015–1022.

Kahn, R. L., Goldfarb, A. I., Pollack, M., & Peck, A. (1960). Brief objective measures for the determination of mental status in the aged. *The American Journal of Psychiatry*, *117*, 326–328.

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., & Jaffe, M. W. (1963). Studies of illness in the aged: The index of ADL: A standardized measure of biological and psychological function. *Jama*, *185*, 914–919.

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist*, *9*, 179–186.

Lezak, M. D., Howieson, D. B., Bigler, E., & Tranel, T. (2012). *Neuropsychological assessment* (5th ed., pp. 260). New York, NY: Oxford University Press.

Lindeboom, J., & Jonker, C. (1989). *Amsterdam Dementie-Screening Test. [Amsterdam Dementia-Screening Test]*. Amsterdam: Pearson Assessment and Information B.V.
Lindeboom, J., & Schmand, B. (2003). *Visual Association Test. Manual*. Amsterdam: Hogrefe Publishing bv.
Lindeboom, J., Schmand, B., Meyer, S. R. A., & De Jonghe, J. F. M. (2014). *Visual Association Test. Manual*. Amsterdam: Hogrefe Publishing bv.
Lindeboom, J., Schmand, B., Tulner, L., Walstra, G., & Jonker, C. (2002). Visual association test to detect early dementia of the Alzheimer type. *Journal of Neurology, Neurosurgery, and Psychiatry*, *73*, 126–133.
Lowndes, G. J., Saling, M. M., Ames, D., Chiu, E., Gonzalez, L. M., & Savage, G. R. (2008). Recall and recognition of verbal paired associates in early Alzheimer's disease. *Journal of the International Neuropsychological Society*, *14*, 591–600.
Luteijn, F., & Barelds, D. P. H. (2004). *GIT 2: Groninger Intelligentie Test 2. [GIT 2: Groninger Intelligence Test 2]*. Amsterdam: Harcourt Assessment B.V.
Meyer, S. R. A., de Jonghe, J. F., Schmand, B., & Ponds, R. W. H. M. (2017). The Visual Association Test-Extended: A cross-sectional study of the performance validity.measures. *The Clinical Neuropsychologist*, *31*, 789–813.
Meyer, S. R. A., & De Jonghe, J. F. M. (2017). *Visuele associatietest – Extended. Handleiding. [Visual Association Test-Extended. Manual]*. Amsterdam: Hogrefe Publishing bv.
Meyer, S. R. A., Spaan, P. E. J., Boelaarts, L., Ponds, R. W. H. M., Schmand, B., & de Jonghe, J. F. M. (2016). Visual associations cued recall A paradigm for measuring episodic memory decline in Alzheimer's disease. *Aging, Neuropsychology, and Cognition*, *23*, 566–577.
Morris, J. C. (1993). The clinical dementia rating (CDR): Current version and scoring rules. *Neurology*, *43*, 2412–2414.
Osterreith, P. A. (1944). *Le test de copie d'une figure complex* [The complex figure copy test]. *The Clinical Neuropsychologist*, *7*, 9–15.
Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, *74*, 241–263.
Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*, 255–287.
Paivio, A., Rogers, T. B., & Smythe, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science*, *11*, 137–138.
Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, *256*, 183–194.
Petersen, R. C., Doody, R., Kurz, A., Mohs, R. C., Morris, J. C., Rabins, P. V., & Winbald, B. (2001). Current concepts in mild cognitive impairment. *Archives of Neurology*, *58*, 1985–1992.
Pfeiffer, E. (1975). A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. *Journal of the American Geriatrics Society*, *23*, 433–441.
PsychCorp. (2009). *Wechsler Memory Scale-Fourth Edition (WMS-IV) technical and interpretive manual*. San Antonio, TX: Pearson.
Reesink, F. E., Lemstra, A. W., van Dijk, K. D., Berendse, H. W., van de Berg, W. D. J., Klein, M., & Van de flier, W. M. (2010). CSF α-synuclein does not discriminate dementia with Lewy bodies from Alzheimer´s disease. *Journal of Alzheimer´S Disease*, *22*, 87–95.
Reitan, R. M., & Wolfson, D. (1985). *The Halstead-Reitan neuropsychological test battery*. Tucson, AZ: Neuropsychological Press.
Rey, A. (1941). *L'examen psychologique dans les cas d'encephalopathie traumatique* [Psychological examination of traumatic encephalopathy]. *Archives De Psychologie*, *28*, 286–340.
Rey, A. (1964). *L´examen Clinique en psychologie [The clinical examination in psychology]*. Paris: Presses Universitaires de France.
Roth, M., Huppert, F. A., Mountjoy, C. Q., & Tym, E. (1998). *The Cambridge examination of mental disorders of the elderly revised: CAMDEX-R*. Cambridge: Cambridge University Press.
Roy, D. S., Arons, A., Mitchell, T. I., Pignatelli, M., Ryan, T. J., & Tonegawa, S. (2016). Memory retrieval by activating engram cells in mouse models of early Alzheimer's disease. *Nature*, *531*, 508–512.

Saan, R. J., & Deelman, B. G. (1986). *15-woorden test A en B* [15-word test A and B]. A. Bouma, J. Mulder, J. Lindeboom, & B. Schmand (Eds.), 2012. *Handboek neuropsychologische diagnostiek [Handbook neuropsychological assessment]* (2nd 267–282). Amsterdam: Pearson Assessment and Information B.V.

Schmand, B., Deelman, B. G., Hooijer, C., Jonker, C., & Lindeboom, J. (1996). *De item-reeks van de cognitieve screening test vergeleken met die van de mini-mental state examination [The items of the cognitive screening test compared to those of the mini-mental state examination]. Tijdschrift voor Gerontologie en Geriatrie* (Vol. *27*, pp. 29–33).

Schmidt, M. (1996). *Rey auditory verbal learning test. A handbook*. Los Angeles: Western Psychological Services.

Sheik, J. I., & Yesavage, J. A. (1986). Geriatric depression scale (GDS): Recent evidence and development of a shorter version. *Clinical Gerontologist*, *5*, 165–173.

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, *6*, 156–163.

Sivan, A. B. (1992). *Benton visual retention test* (5th ed.). San Antonio, TX: PsychoCorp/Pearson.

Tounsi, H., Deweer, B., Ergis, A., Van der linden, M., Pillon, B., Michon, A., & Dubois, B. (1999). Sensitivity to semantic cueing: An index of episodic memory dysfunction in early Alzheimer disease. *Alzheimer Disease and Associated Disorders*, *13*, 38–46.

Tulving, E., & Madigan, S. A. (1970). Memory and verbal learning. *Annual Review of Psychology*, *21*, 437–484.

Ulrich, R., Nitschke, J., & Rammsayer, T. (2006). Perceived duration of expected and unexpected stimuli. *Psychological Research*, *70*, 77–87.

UNESCO. (2006). *International Standard Classification of Education: ISCED 1997 (re-edition)*. Montreal: Author.

Van der Vlies, A. E., Verwey, N. A., Bouwman, F. H., Blankenstein, M. A., Klein, M., Scheltens, P., & van der Flier, W. M. (2009). CSF biomarkers in relationship to cognitive profiles in Alzheimer disease. *Neurology*, *72*, 1056–1061.

Van Toutert, M., Diesfeldt, H., & Hoek, D. (2016). De cognitieve screening test (cst) bij normale cognitieve veroudering en dementie: Drie varianten en bruikbaarheid voor de klinische praktijk [The cognitive screening test (CST) for normal cognitive functioning and dementia: Three subsets and usefulness for clinical practice]. *Tijdschrift Voor Neuropsychologie*, *3*, 274–292.

Verhage, F. (1964). Intelligentie en leeftijd: Onderzoek bij Nederlanders van twaalf tot zevenenzeventig jaar [Intelligence and age: Research within Dutch participants aged from 12 to 77 years old]. Dissertation. Assen: Van Gorcum.

Wechsler, D. (1945). A standardized memory scale for clinical use. *Journal of Psychology*, *19*, 87–95.

Wilson, B. A., Alderman, N., Burgess, P., Emslie, H., & Evans, J. J. (1996). *Behavioural assessment of the dysexecutive syndrome (BADS)*. Bury St. Edmunds, U.K.: Thames Valley Test Company.