The Hopkins Verbal Learning Test and Detection of MCI and Mild Dementia: A Literature Review

Xin Xu[1,4*], Tri Budi Rahardjo[2], Shifu Xiao[3], Eef Hogervorst[1#]

[1]Psychology Division, School of Sport, Exercise and Health Sciences, Loughborough University, the United Kingdom, UK

[2]Center for Aging Studies, Depok, Indonesia

[3]Shanghai Mental Health Center, Shanghai, China.

[4]Memory, Ageing and Cognition Center, National University of Singapore

[*]**Corresponding author**: Xin Xu, Applied Cognitive Research, Psychology Division, School of Sport, Exercise and Health, Brockington Building, AshbyAsby Road, Loughborough University, Loughborough LE11 3TU, UK Tel: + 44 1509 223020, E-mail: X.Xu@lboro.ac.uk

[#]**Corresponding author**: Eef Hogervorst, Applied Cognitive Research, Psychology Division, School of Sport, Exercise and Health, Brockington Building, AshbyAsby Road, Loughborough University, Loughborough LE11 3TU, UK Tel: + 44 1509 223020, E-mail: E.Hogervorst@lboro.ac.uk

Abstract

A decline in verbal memory is one of the primary symptoms found in amnestic Mild Cognitive Impairment (aMCI) and most types of dementia. To detect verbal memory impairment in clinical and research settings, the Hopkins Verbal Learning Test (HVLT) has been recommended. In the current paper, the utility of the HVLT for aMCI/mild dementia diagnostics, as well as its use in treatment studies is reviewed. The HVLT was considered to have good diagnostic accuracy, be well tolerated and applicable across cultures. Optimal cut-offs for MCI and mild dementia are discussed and their possible relation to demographic factors such as age, gender, ethnicity and education.

Keywords: Hopkins verbal learning test; Mild cognitive impairment; Dementia

Background

With an advancing age, there is an average decline in various areas of cognitive function, such as episodic memory and speed of complex information processing [1]. Dementia is a separate progressive neurodegenerative disorder that causes a severe decline in memory and other cognitive abilities, which have a significant impact on the quality of life [2,3]. There is currently no effective treatment. Globally, the number of people afflicted with dementia has shown a steady growth over the last decades [2,3]. Mild cognitive impairment (MCI) is defined as cognitive decline worse than that of those who have a similar advanced age, but which causes no interference with activities of daily life, such as dementia does. The most commonly used MCI diagnostic criteria were developed by Petersen et al. and confer an increased risk for dementia [4]. It has been reported that 10-12% of individuals with MCI progress to dementia per year [5]. There is a growing awareness of MCI, where many studies now focus on the discrimination between those undergoing the normal cognitive aging process and those with MCI, who may convert to dementia. It may be that future interventions have a better chance of success in those who have not developed dementia yet, but are at risk for this. Good screening methods for MCI and early dementia are imperative. In this paper we review the Hopkins Verbal Learning Test (HVLT) and its ability to discriminate between people with mild dementia and MCI as compared to non-afflicted controls.

The HVLT [6] is a word-learning test measuring episodic verbal memory. Version A consists of 12 words from 3 low frequency categories (human shelter, animals and precious stones), which are also late acquired words during development. These words are read out loud after which the participant recalls them in any order. 20 -30minutes after obtaining the total immediate recall (reflecting learning ability, which is obtained by repeating the same word list 3 times and adding up all correctly recalled words over the 3 trials), a delayed recall without cues or prompting is done. The HVLT should be particularly adept at identifying people with amnestic MCI (aMCI), where according to Petersen [5] the primary distinction between control subjects and subjects with aMCI is in the area of verbal memory. To reduce slight learning effects in controls, six parallel versions exist, which have shown good inter-test reliability [6]. The HVLT has been shown to have good validity and reliability and is well tolerated by elderly people [7].

Here we review papers investigating the discriminative capacity of the HVLT to identify patients with MCI versus controls to establish whether similar cut-offs of the total immediate recall for screening were identified among different studies. We also included papers investigating participants with mild dementia, as the distinction between MCI and mild dementia is often not entirely clear.

Methods

Data Sources

The PUBMED electronic database was systematically scanned using different combinations of search terms. There was no restriction on year of publication. The references of the included studies were searched for relevant articles (n=5). The last search was performed on the 8th of July, 2013. Using the term 'Hopkins Verbal Learning Test', 220 relevant publications were found. Using a combination of 'Hopkins Verbal Learning Test', 'dementia', 46 results were found. A combination of 'Hopkins Verbal Learning Test' and 'MCI', only rendered 13 relevant publications. After screening by reviewing abstracts, 26 articles were included for the full literature review. The schema below describes this process in more detail (Figure 1).

Figure 1: Selection process

Results

The utility of the HVLT as a screening test for MCI and dementia

We first investigated the most optimal cut-off scores when screening for MCI and mild dementia versus controls using the HVLT total immediate recall. According to a study by Hogervorst et al [8] in an Oxfordshire (UK) based cohort of carefully matched cases and controls, 87% sensitivity and 98% specificity for mild to moderate dementia (versus controls) was obtained using a cut-off score of 14/15 of the HVLT total recall, whereas a cut-off score of 18/19 yielded better sensitivity (95%), but somewhat lower specificity (77%). Similarly, for mild dementia in Australia [9], the HVLT total immediate recall had a sensitivity of 96% and a specificity of 80% using the same cut-off score

of 18/19. Significantly different HVLT total recall scores between age and education equated controls, patients with MCI, and with cerebrovascular disease (CVD, which included vascular cognitive impairment and vascular dementias) as well as Alzheimer's disease (AD, the most common form of dementia) were also found in another Oxfordshire case-control study (on average: 26, 18, 17 and 10 words, respectively, were recalled per group)[10]. In addition, in this study when using a cut-off score of 21.5 for the total recall, 78% sensitivity and 80% specificity was reported at baseline between 51 healthy controls and 15 control participants who would develop MCI after a 2-3 year follow-up. A third Oxfordshire study [11] gave similar data on specificity and sensitivity for AD, MCI and controls as found in the previous Oxford based studies, and these were again maintained at follow-up. All those with dementia declined, controls all improved and half of MCI showed a decline in function, similar to the dementia cases (see table 1& 2 for the studies described above). From these studies, which had al matched or equated for age, gender ratio and education, it may be suggested that a HVLT total immediate recall cut-off score of around 14/15 for dementia overall, and below 18/19 for mild dementia is best used for screening. Table 2 suggests that for MCI vs. controls, a cut-off score of 24/25 probably gives best sensitivity (with around 22 for Chinese populations). Between MCI and AD, the best cut-off score is around 16/17 word recalled on the total immediate recall.

Study	Setting	Sample Size			Mean Age				Gender (Female %)	Education (years)	Ethnicity
		NCI	MCI	Dementia	NCI	MCI	Dementia	Total			
Barandt [6]		18	3 Amnesic	45 AD	73.6	77	69.4				
Benedict [12]	Community and others	541						48.1	63.0%	13.8	85% caucasian, 14% African, 1% others
Lacritz [13]		25			70.7					16.2	
Shapiro [7]		59		55 AD	75.3	76.1	76.1			11.5	
		37		34 VaD	73.6	73.1	73.1			11.4	
Frank [9]	Clinic	15		26				74.7	62.5%	8.5	Caucasian
Lacritz [14]				40 AD		73.4				13.7	
Hogervorst [8]	Community	114		82				75	44.4%	11	Caucasian
Friedman [15]	Community	237						range from 60-84		n=173 <= 12 years; n=42 12 years; n=22 > 12 years	African American
De Jager [10]	Community	51	29	60 AD / 12 CVD	75	76	AD:77 CVD:75		50.7%	12	
Kuslansky [16]	Community	323		70	78.6		82		36.4%	12.3	mixed (61.8% Caucasian)
Hester [17]	Community	203			73.1				55.2%	11.1	

Schrijnemaekers [11]	Community	54	19	28	76	76.2	77.2	76.4	48.5%	18.2	Caucasian
Gaines [18]	Clinic	30		61	63.5		72.6		62.3%	12.4	mixed (75.8% Caucasian)
Cherner [19]	Community	125 (normative sample)						37.5	42.5%	9.7	Mexican decent
De Jager [20]	Clinic	98	21			77.2	82				
Foster [21]	Clinic	31		19 mod AD; 9 mild AD	65.7		74.7			14	
Aretouli [22]	Clinic							63.6	49.4%	14.7	
Baek [23]	Clinic		112	97 AD		69.1	73.5			11.1	
Gómez-Tortosa [24]	Cohort		103 with DI <8; 107 with DI>=8					72.3	64.3%	72.4% below level 2 (out of 5)	
Shi [25]	Clinic	249	134	97	67.1	69.9	71.1		58.5%	11	
McLaughlin [26]	Clinic			22 DLB; 32 AD			DLB:77.9 AD:78.9		66.7%	12.5	
Gonza´lez-Palau [27]	Community + Clinic	109	132	54	81.9	82	83.4		74.8%	8.3	

Table 1: Demographics of the reviewed studies

Notes: AD = Alzheimer's Disease; VaD = Vascular Dementia; CvD = Cerebrovascular Disease; DI = Discrimination Index; mod AD = Moderate AD; DLB = Dementia with Lewy Bodies

Study	Aim of Comparison	HVLT total recall performance			
		AUC (95% CI)	Optimal Cut-off	Sensitivity	Specificity
Brandt et al. [6]	NCI vs Amnesic & Dementia		19/20	0.94	1
Frank et al. [9]	NCI VS mild Dementia	-	18	0.96	0.8
Hogervorst et al. [8]	NCI VS Dementia	0.97	14.5	0.87	0.98
		(0.95-0.99)	19.5	0.95	0.77
De Jager et al. [10]	NCI VS MCI VS AD VS CVD	0.88	15.5 (MCI VS AD)	0.91	0.69
		0.84	14.5 (CVD VS AD)	0.82	0.75
Kuslansky et al. [16]	NCI VS Dementia	0.89	<16	0.83	0.83
Schrijnemaekers et al. [11]	NCI VS MCI VS AD		24.5 (NCI VS MCI)	0.82	0.79
			16.5 (MCI VS AD)	0.79	0.96
De Jager et al. [20]	NCI VS MCI	0.9	25.5	0.79	0.95
Shi et al. [25]	NCI VS MCI VS Dementia	0.98	15.5 (NCI vs Dementia)	0.95	0.93
		0.79	21.5 (NCI vs MCI)	0.69	0.71

		0.95			
		(0.92-0.98)	13	0.96	0.85
Gonza´lez-Palau et al. [27]	NCI VS AD	0.77			
		(0.65-0.89)	12.5 (AD VS HD)	0.97	0.52
		0.64			
		(0.48-0.80)	12.5 (AD VS PD)	0.72	0.52
		0.66			
Aretouli et al. [22]	AD VS HD VS PD	(0.49-0.83)	13.5 (PD VS HD)	0.87	0.44

Table 2: The discriminate ability of the HVLT in differentiating between diagnostic groups in the reviewed studies which reported AUC and optimal cut-off scores Notes: NCI=No Cognitive Impairment; MCI=Mild Cognitive Impairment; AD=Alzheimer' Disease; CVD= Cerebral Vascular Disease; HD= Huntington's disease; PD=Parkinson's disease.

HVLT assessment at baseline as part of treatment trials for those with MCI

In this section we investigated the ability of the HVLT to detect treatment effects. The HVLT is regarded not only as a good test for the screening and detection of memory impairment, but also to asses treatment effects in participants with MCI. For instance, as individuals with MCI could benefit from learning strategies during word recall tasks; this could be further enhanced by pharmacological and-non pharmacological treatments [28]. The HVLT was used successfully to assess effects of cognition enhancers in elderly both without dementia [29] and with dementia, which included treatment using Chinese herbal medicine [30]. HVLT has also been used to test cognitive improvement using other non-pharmacological techniques in elderly [31].These results demonstrate the applicability of verbal memory tests, such as the HVLT, in determining effective treatments for normal cognitive ageing, as well as the mild decline in cognitive ability in MCI and the more severe decline in dementia. The HVLT could be used for all these groups and thus for multiple purposes, such as both for baseline screening purposes, as well as subsequent treatment trials, which would be cost-effective.

Demographic factors and HVLT performance

In this section, we describe potential limitations in the use of fixed cut-off scores for the HVLT, if these performance scores are affected by demographic factors. In our matched case control studies (see above) analyses suggested this was not the case. However, many of the other studies investigating the HVLT (see table 2) had not matched or equated cases and controls for age, gender ratio, depression and education, which can all affect performance, and often MCI had not been included with controls (see Hogervorst for a discussion [8])For instance, in the first paper describing the HVLT [6]cases and controls were not comparable in demographic factors and systematic differences between groups (in age, gender ratio, education etc.) could be responsible for the very large differences reported. This could also explain some of the differences in reported optimal cut-off scores [9-11,16,20,22,25,27].

Demographic factors could also potentially affect finding treatment effects, especially if both treatment and test scores are affected by these (e.g. gender affecting verbal memory performance differences, and hormones differentially affecting genders in verbal memory

performance [32]). Although, for instance, Kuslansky [16] did not find any age, sex and education differences on single HVLT test performance in a multi-ethnic cohort, a number of other authors reported a significant influence of demographic factors on HVLT recall performance. In one study [20], age was identified as the best predictor of the HVLT total immediate recall score, when age, years of further education, gender, activities of daily living (ADL) and subjective memory complaints (SMC) were entered as independent variables in regression analyses.

However, in this study the MCI group was significantly older (81.95±5.4 VS 77.18±5.9, p=0.001) and had more SMCs (1.76±1.04 VS 1.32±1.00, p=0.032 based on a 0-4 range report) than the control group, so these results may have been susceptible to systematic confounds. Despite these possible confounds, a cut-off score of 25.5 on the total immediate recall score of the HVLT (similar to the other better matched studies mentioned above) rendered 79% sensitivity and 95% specificity when discriminating between controls and those with MCI. Cherner [19] examined a sample of middle-aged Spanish speakers with an average low educational level from the U.S.–Mexico border region on their HVLT-R performance and found that education (p<0.001), rather than age or gender (24.94± 4.47 for males and 25.44±4.29 for females), was significantly related to the HVLT-R total immediate recall score. Age in this study was only found to be related to the Recognition Discrimination Index of the HVLT-R, which we do not use, as it did not add to diagnostic discrimination [8]. Friedman [15] reported significant, but moderate-sized effects of education in a community based African-American sample, as well as effects of age and gender (p<0.01) on the HVLT-R test performance. Another Australian study reported a significant impact of age and education, but not gender, on the HVLT-R total immediate recall [17].

Gender differences were thus found in several(but not all) studies, as women are often reported to have better performance on tests involving verbal components, while men are thought to perform better on tests involving visuo-spatial skills [33-35]. It becomes more difficult when cultural differences seem to further modify these demographic factors, such as gender. For instance, we found significant differences by gender when predicting HVLT total immediate recall performance in China [36], but no overall effects of gender on HVLT total immediate recall scores in Indonesia, although gender differences were also seen in some ethnic groups here, but in the opposite direction to those found in China [37]. Aging seemed to further modify the gender

effect by culture [12]. However, in a meta-analysis, this age by gender differences could be largely explained by systematic differences in health status and prior education obtained between genders [38]. In table 3, we reported the immediate recall (IR) and delayed recall (DR) performance on the HVLT based on 3 reviewed multi-ethnic studies which provided normative data for the HVLT and which were stratified by age and education [15,17,12].

Study	Stratification							HVLT,mean (SD)	
	Age, years		Education, years		Gender, n				
	Mean (SD)	Range	Mean (SD)	Range	Male	Female	Sample size	IR	DR
Benedict [13]	24.2 (4.6)	17-30	13.8 (2.1)	8 - 18	46	56	102	29.4 (3.7)	10.6 (1.6)
	42.1 (6.5)	31-54	13.8 (1.9)	10 - 20	79	156	235	28.8 (3.8)	10.3 (1.7)
	61.9 (4.3)	55-69	13.8 (2.6)	6 - 20	50	79	129	27.5 (4.3)	9.8 (1.8)
	75.2 (4.5)	70-88	13.4 (2.9)	5 - 20	25	50	75	25.2 (5.5)	8.7 (2.8)
Friedman [19]		60-71		<12	37	30	67	16.9 (3.2)	6.5 (1.5)
				12	11	16	27	18.6 (2.5)	6.7 (1.4)
				>12	6	11	17	20.5 (4.7)	6.8 (2.5)
		72-84		<12	49	57	106	14.9 (4.2)	5.6 (2.0)
				12	4	11	15	18.5 (3.6)	7.0 (2.0)
				>12	1	4	5	17.8 (5.7)	5.2 (3.1)
Hester [15]		60-69		<=10			29	20.0 (5.5)	6.3 (3.3)
				>10			35	2.6 (4.8)	8.4 (2.9)
		70-79		<=10			63	19.4 (5.8)	6.4 (3.5)
				>10			45	20.2 (4.6)	7.3 (2.7)
		80-89		<=10			15	17.4 (5.2)	5.4 (3.1)
				>10			16	21.1 (4.6)	5.4 (2.6)

Table 3: Normative data for the HVLT IR and DR performance using age and education strata

From this table it seems clear that less education decreases performance, and there may also be age and ethnic differences. However, as stated differences in health status (which can be affected by age) rather than these factors per se may affect performance and this was not controlled for.

The ethnic/cultural confounds are also not specific to the HVLT. Hogervorst [32] and Schwartz [39] reported disparities in memory recall of other word lists by ethnicity, even when controlling for age, sex and education. On the other hand, Tanaka [38] failed to find better verbal learning on another (California) verbal learning test (CVLT) in a European American group when compared to that of Japanese Americans, who outperformed them. These differences were hypothesized to be related to an inherent systematic bias, with ethnic/cultural differences in educational quality and different cultural exposures to learning and vocabulary.

On the other hand, despite having slightly lower average HVLT total immediate recall scores, less educated Indonesian rural participants still had the same cut-off scores for dementia as the highly educated Oxfordshire cohorts [37]. This would argue against the need for educationally and ethnicity specific cut-off scores.

The Chinese version of the HVLT was administered to differentiate between aMCI, dementia (subsequently divided into AD and all types of dementia) and controls [25]. There was a wide range of performance on the HVLT total immediate recall between groups (with on average 23.8 words recalled for controls, 18.0 for aMCI, 6.1 for AD, and 6.4 for all types of dementia, p<0.001). However, after applying an age split (50-64 vs 65-80 age group), a more optimal cut-off of 18.5 was obtained with 96% sensitivity and 92% specificity, when distinguishing AD from controls in the younger group (50-64), whereas a cut-off of 14.5 was found to be more accurate for the older group (65-80) with a sensitivity of 95% and specificity of 93%. This was similar to the results from the Oxfordshire and Indonesian data which participants were within this age-range. On the other hand, in those with early onset AD (before age 65 years) also assessed in Leicester (UK), an optimal cut-off score of 19 was found with 100% sensitivity and specificity (Clifford, in press), also similar to the findings from China as mentioned above [25]. Our data from older Chinese institutionalized elderly with an average age of 80 years also suggested that a lower cut-off of 10/11 words on the HVLT total immediate recall should be used for dementia, particularly when comparing elderly with psychiatric disorders to those with dementia [36]. These data taken together would suggest that regardless of

culture and education, when adequate back-translation and adjustment has been done, age-related cut-offs (<65 and >80 years of age) may still yield better specificity and sensitivity for dementia screening.

Importantly, an older age, the female gender and low education are all risk factors for dementia [2], so many case control cohorts (unless matching was done) will have these systematic biases. These systematic differences are difficult to control for in analyses, as they are inherent to being a case and not a control. In addition, systematic differences in exposure to a particular vocabulary perhaps or coping skills which can aid learning skills are difficult to control for, even when education is controlled for [32] For these reasons we always translated and back-translated and if items were not recognized, adapted the list conform local knowledge (e.g. but these items would still be within the category of semi-precious stones, animals or human shelter). As stated, when case-control cohorts have been carefully matched for education, gender and age, these demographic variables, however, do not contribute to differences in HVLT total immediate recall performance associated with dementia and the cut-off scores[8]. In addition, as mentioned above, even when cases and controls are not matched for age, gender or education, our Indonesian and Chinese data (when compared to UK data) suggested that cut-off scores for MCI and mild dementia may be remarkably similar across cultures, except in patients with early onset AD (age < 65 years) or the institutionalized oldest old (>80).

validated the HVLT in different language (e.g. using English, Spanish, Chinese, Indonesian and Korean) versions 6.8 [32][38][27][40][41]. This has not always led to controls obtaining similar scores on the total immediate recall, as would be expected (see an earlier discussion also on Indian and Indonesian controls scoring lower on the Mini Mental State Examination (MMSE) and the HVLT when compared to age equated elderly in the UK in [35]). For instance, the Korean version of the HVLT (K-HVLT) was investigated in Korean MCI and AD patients [23]. The total score of this HVLT showed correlations with the MMSE, Clinical Dementia Rating (CDR), Global Deterioration Scale (GDS), and Story Recall Test (SRT), but also revealed significantly different levels of memory performance in MCI and AD patients compared to a control group (mean score 20.6 for controls; 16.3 for MCI, and 12.4 for AD, $p<0.001$). Similar to the studies mentioned in [37], controls reached a lower average score on the total immediate recall when compared to other control cohorts mentioned earlier (also see table 2). However, average scores for the MCI and AD cases were quite similar to the earlier mentioned studies (see above).

Similarly, the Spanish version of the HVLT total recall [24] also showed different levels of performance among different diagnostic groups (11.7 for MCI, 9.63 for AD and 17.7 for controls). However, here average HVLT total immediate recall scores for controls were even3-4 points lower than those from the Korean study (so around 6-7 points lower than those of other studies investigating controls in table 4).

Different Language Versions of the HVLT in Detecting Cognitive Impairment

The HVLT has been widely translated and used in different countries, but there are only a limited numbers of studies which

Study	Aim of Comparison	Comparing Group	HVLT performance		Significance
			IR	**DR**	
Lacritz et al. [12]	HVLT-R VS CVLT	HVLT	26.3 (4.9)	10.2 (1.8)	IR: r=0.74, p<0.001; DR: r=0.65, p<0.001
		CVLT	50.2 (9.7)	11.9 (2.5)	
Shapiro et al. [7]	NCI VS AD	NCI	24.8 (5.1)	-	F=164.8, p<0.001
		AD	12.2 (5.3)	-	
	NCI VS VaD	NCI	24.6 (5.5)	-	F=56.4, P<0.001
		VaD	14.5 (5.8)	-	
Lacritz et al. [14]	HVLT-R VS CVLT	HVLT	11.4 (3.5)	0.6 (1.3)	IR: r=0.36, p=0.02; DR: r=0.62, p<0.001
		CVLT	19.2 (6.9)	1.2 (2.0)	
Gaines et al. [41]	NCI VS Dementia	NCI	25.3 (4.9)	8.9 (2.1)	p<0.05
		VaD	15.1 (4.8)	2.9 (2.5)	
		AD	12.2 (4.8)	1.3 (1.5)	
Cherner [15]	Demographic correction	Age, education , and gender difference	25.2 (4.3)	8.4 (2.4)	IR: r=0.36, p<0.001 (education); DR: r=0.43, p<0.001 (education); NS for age and gender

Foster et al. [24]	NCI VS Mild AD VS mod AD	NCI	26.0 (9.6)	4.4 (2.2)	IR: F=54.47, p<0.0001
		mild AD	14.0 (1.5)	4.7 (2.5)	
		moderate AD	8.8 (1.0)	1.0 (2.1)	
Baek et al. [18]	NCI VS MCI VS AD	NCI	20.6 (4.3)	6.4 (2.1)	IR: F=66.12, p<0.001; DR: F=81.6, P<0.001
		MCI	16.3 (4.3)	3.4 (2.5)	
		AD	12.4 (4.1)	1.4 (2.1)	
McLaughlin et al. [26]	DLB VS AD	DLB	2.5 (1.8)	1.2 (2.0)	IR: p<0.04; DR: p=0.01
		AD	1.6 (1.1)	0.0 (0.3)	
Gómez-Tortosa et al. [13]	Baseline	MCI (DI<8)	15.4 (3.0)	2.3 (1.9)	IR: p<0.001; DR: p<0.0001
		MCI (DI>=8)	16.8 (3.0)	3.5 (1.9)	
	48 ±12 Months follow-up	MCI (stable)	16.7 (3.0)	3.6 (1.8)	IR: p=0.001; DR: p<0.0001
		MCI (progression)	15.2 (3.1)	1.8 (1.9)	

Table 4: Comparison of the HVLT immediate and delay recall performance. Notes: CVLT=California Verbal Learning Test; IR=Immediate Recall; DR= Delayed Recall; VaD= Vascular Dementia; AD= Alzheimer' Disease; DLB= Dementia with Lewy Bodies; DI= Discrimination Index based on HVLT performance

Many participants in this study had low levels of education (72%). This perhaps further reflected the much lower cut-off scores required per diagnostic group in that study. An HVLT total immediate recall cut-off score of <=14 words demonstrated a 70.1% sensitivity and 73.7% specificity when discriminating aMCI cases from controls, and a score of <=11 words recalled had a 79.2% sensitivity and 91.9% specificity, when differentiating AD cases from controls. Slight differences in age and/or education between cases and controls between and within cohorts as mentioned above may result in different results reported. Alternatively, this lower cut-off score for optimal specificity could perhaps suggest inadequate adaptation of the word lists to local knowledge. As stated in the previous paragraph, adequate back translation and adaptation to local knowledge may go some way in solving these issues.

Comparing the HVLT with Other Memory Tests

Several perhaps more commonly used verbal learning tests were compared to the HVLT. For instance, the California Verbal Learning Test (CVLT, mentioned earlier) is also used in the screening of cognitive impairment and it showed strong correlations with the HVLT [14]. While the authors summarized that the HVLT may not always be challenging enough, nevertheless, the HVLT was suggested to be a superior multidimensional brief verbal learning assessment when compared to the CVLT, as it took less time and training to use. In addition, the CVLT as discussed before also is affected by demographic factors. For instance, in Norman's study [42], age, education, ethnicity, and gender were also found to be significant predictors of CVLT total recall performance among both Caucasians and Africa American populations. Another comparative study between the HVLT and the SRT [23] also concluded that although the SRT was well correlated with the HVLT, the HVLT was less influenced by education and would thus also be deemed superior.

Discussion

In conclusion, the HVLT has been shown to be an effective instrument in the screening of MCI and mild dementia with a high level of sensitivity and specificity. Furthermore, the HVLT could play a role in treatment studies in MCI and mild dementia patients as its baseline screening assessment could be included in the assessment and save money and time.

However, the effect of demographic factors on verbal memory performance remains to be debated. Ethnic differences reported in this review [15][32][39] were probably confounded by systematic differences in exposure to vocabulary, health status and educational levels. Despite these differences, Chinese and rural Indonesian elderly did not require specific cut-off scores for MCI and mild dementia when compared to our UK cohorts, but some of the words in our studies had been changed to fit regional knowledge. To control for some of the possible health status effects, age adjusted norms for the HVLT may be important for early onset AD (<65), as well as for those with advanced age (>80) [43]. As discussed, the Chinese version of the HVLT using age strata resulted in a 4 points difference on the cut-off score (18.4 VS 14.5) to obtain maximum discriminative capacity for these groups[25].

The HVLT not only has the ability to differentiate MCI from controls, but can also distinguish between different stages of cognitive impairment which is useful in treatment and diagnostic assessments. A revised and copyrighted version of the HVLT (-R) added a DR and delayed recognition trial which demonstrated good reliability [12]. The total IR trials were published in the public domain [6] and so whether these are copyrighted could be debated. One study showed significantly worse HVLT IR mean scores for the MCI progression group when compared to the MCI stable group (15.2 vs 16.7, p=0.001), whereas an even greater difference was found between groups on the HVLT DR mean scores (1.8 vs 3.6, p<0.0001)[24].Others reported that DR was less susceptible to educational confounding effects in analyses of another similar word learning test [40].On the other hand, studies

reported that the HVLT total immediate recall score proved to be useful in discriminating moderate to severe AD and mild AD from controls (with average scores of 8.8 for moderate to severe AD, 14.0 for mild AD and 26.0 for controls)without being confounded by educational effects [21]and using the total immediate recall only would save time in assessments. However, as no significant difference between mild AD and moderate to severe AD in IR scores was detected [21] perhaps other tests and including the DR should be used to further discriminate between these stages.

Lastly, the HVLT was found to distinguish between dementia with AD from that with Lewy Bodies (DLB) [26], and perhaps Parkinson's (PD) and Huntington's (HD) disease [20]. The HVLT showed its superiority in distinguishing those with executive dysfunction from healthy controls with no executive impairment (16.9 vs 15.0, p=0.02) [41]. This is perhaps reflective of its ability to also distinguish between CVD (vascular cognitive impairment/dementia, such as VCI/VaD) versus AD and controls as mentioned previously [7,8,18]. VCI/VaD usually present with executive function impairment before memory problems becomes evident. These findings elucidated an even wider range of utility of the HVLT in future studies.

In sum, the HVLT is a useful test which is available in the public domain and which may have a wide range of applicability, from diagnosing MCI, CVD, DLB and AD, to tracking treatment effects.

References

1. Hogervorst E (1998) Age-related decline and cognition enhancers. Neuropsych publishers: Maastricht.

2. Alzheimer's Association (2010) Alzheimer's disease facts and figures. Alzheimers Dement 6: 158-194.

3. Alzheimer's Association (2014) Alzheimer's disease facts and figures. Alzheimers Dement 10: e47-92.

4. Petersen RC (2004) Mild cognitive impairment as a diagnostic entity. J Intern Med 256: 183-194.

5. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, et al. (1999) Mild cognitive impairment: clinical characterization and outcome. Arch Neurol 56: 303-308.

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

7. Shapiro AM, Benedict RH, Schretlen D, Brandt J (1999) Construct and concurrent validity of the Hopkins Verbal Learning Test-revised. Clin Neuropsychol 13: 348-358.

8. Hogervorst E, Combrinck M, Lapuerta P, Rue J, Swales K, et al. (2002) The Hopkins Verbal Learning Test and screening for dementia. Dement Geriatr Cogn Disord 13: 13-20.

9. Frank RM, Byrne GJ (2000) The clinical utility of the Hopkins Verbal Learning Test as a screening test for mild dementia. Int J Geriatr Psychiatry 15: 317-324.

10. De Jager CA, Hogervorst E, Combrinck M & Budge MM (2003) Sensitivity and specificity of neuropsychological tests for mild cognitive impairment, vascular cognitive impairment and Alzheimer's disease. Psychological Medicine, 33:1039-1050.

11. Schrijnemaekers A MC, De Jager CA; Hogervorst E & Budge MM (2006) Cases with Mild Cognitive Impairment and Alzheimer's Disease Fail to Benefit from Repeated Exposure to Episodic Memory Tests as Compared with Controls. Journal of Clinical and Experimental Neuropsychology, 28:438 – 455.

12. Benedict RHB & Zgalijardic DJ (1998) Practice effects during repeated administration of memory tests with and without alternate forms. Journal of Clinical and Experimental Neuropsychology, 20: 339-352.

13. Lacritz LH, Cullum CM (1998) The Hopkins Verbal Learning Test and CVLT: a preliminary comparison. Arch Clin Neuropsychol 13: 623-628.

14. Lacritz LH1, Cullum CM, Weiner MF, Rosenberg RN (2001) Comparison of the hopkins verbal learning test-revised to the California verbal learning test in Alzheimer's disease. Appl Neuropsychol 8: 180-184.

15. Friedman MA, Schinka JA, Mortimer JA, Graves AB (2002) Hopkins Verbal Learning Test-Revised: norms for elderly African Americans. Clin Neuropsychol 16: 356-372.

16. Kuslansky G, Katz M, Verghese J, Hall CB, Lapuerta P, et al. (2004) Detecting dementia with the Hopkins Verbal Learning Test and the Mini-Mental State Examination. Arch Clin Neuropsychol 19: 89-104.

17. Hester RL, Kinsella GJ, Ong B & Turner M (2004) Hopkins Verbal Learning Test: Normative data for older Australian adults. Australian Psychologist, 39: 251 – 255

18. Gaines JJ, Shapiro A, Alt M, Benedict RH (2006) Semantic clustering indexes for the Hopkins Verbal Learning Test-Revised: initial exploration in elder control and dementia groups. Appl Neuropsychol 13: 213-222.

19. Cherner M, Suarez P, Lazzaretto D, Fortuny LA, Mindt MR, et al. (2007) Demographically corrected norms for the Brief Visuospatial Memory Test-revised and Hopkins Verbal Learning Test-revised in monolingual Spanish speakers from the U.S.-Mexico border region. Arch Clin Neuropsychol 22: 343-353.

20. De Jager CA, Schrijnemaekers AC, Honey TE, Budge MM (2009) Detection of MCI in the clinic: evaluation of the sensitivity and specificity of a computerised test battery, the Hopkins Verbal Learning Test and the MMSE. Age Ageing 38: 455-460.

21. Foster PS, Valeria D, Crucian GP, Rhodes RD, Shenal BD &Heilman KM (2009) Verbal learning in Alzheimer's disease: Cumulative wordknowledge gains across learning trials. Journal of the International Neuropsychological Society, 15: 730 – 739

22. Aretouli E, Brandt J (2010) Episodic memory in dementia: Characteristics of new learning that differentiate Alzheimer's, Huntington's, and Parkinson's diseases. Arch Clin Neuropsychol 25: 396-409.

23. Baek MJ, Hyun JK , Hui JR , Seoung HL , Seol HH , Hae RN , Young HC, Chey JY & Kim SY (2011) The usefulness of the story recall test in patients with mild cognitive impairment and Alzheimer's disease, Aging, Neuropsychology,andCognition: A Journal on Normal and Dysfunctional Development, 18: 214-229.

24. Gómez-Tortosa E, Mahillo-Fernández I, Guerrero R, Montoya J, Alonso A, et al. (2012) Outcome of mild cognitive impairment comparing early memory profiles. Am J Geriatr Psychiatry 20: 827-835.

25. Shi J, Tian J, Wei M, Miao Y, Wang Y (2012) The utility of the Hopkins Verbal Learning Test (Chinese version) for screening dementia and mild cognitive impairment in a Chinese population. BMC Neurol 12: 136.

26. McLaughlin NC, Chang AC, Malloy P (2012) Verbal and nonverbal learning and recall in dementia with lewy bodies and Alzheimer's disease. Appl Neuropsychol Adult 19: 86-89.

27. Gonza´lez-Palau F, Franco M, Jime´nez F, Parra E, Bernate M &Solis A (2013) Clinical Utility of the Hopkins Verbal Test-Revised for Detecting Alzheimer's Disease and Mild Cognitive Impairment in Spanish Population. Archives of Clinical Neuropsychology, 28:245–253

28. Ribeiro F, Guerreiro M, De Mendonça A (2007) Verbal learning and memory deficits in Mild Cognitive Impairment. J Clin Exp Neuropsychol 29: 187-197.

29. Yasar A, Zhou J, Varadhan R& Carlson MC (2008) The Use of Angiotensin-converting Enzyme Inhibitors and Diuretics Is Associated With a Reduced Incidence of Impairment on Cognition in Elderly Women. Clinical Pharmacology & Therapeutics,84: 119–126.

30. Lu Y, Zhao W (2001) Clinical study on Chinese herbal, Heart-Regulating Formula and Kidney-Tonifying Formula in improving cognitive ability in Alzheimer's disease. Modern Rehabilitation, 11: 46-47.

31. Tracy J, Ahmed N, Khan W & Sperling MR (2007) A test of the efficacy of the MC Square device for improving verbal memory, learning and attention. International Journal of Learning Technology, 3: 183-202.

32. Hogervorst E, Rahardjo TB, Jolles J, Brayne C & Henderson VW (2011) Gender differences in verbal learning in older participants. Aging Health, 8 : 493-507.

33. Proust-Lima C, Amieva H, Letenneur L, Orgogozo JM, Jacqmin-Gadda H, et al. (2008) Gender and education impact on brain aging: a general cognitive factor approach. Psychol Aging 23: 608-620.

34. Herlitz A, Nilsson LG, Bäckman L (1997) Gender differences in episodic memory. Mem Cognit 25: 801-811.

35. Trenerry RM, Clifford JR, Cascino GD, Sharbrough FW &Ivnik RJ (1995) Gender differences in post-temporal lobectomy verbal memory and relationships between MRI hippocampal volumes and preoperative verbal memory. Epilepsy Research, 20: 69-76.

36. Xu X, Dai J, Hogervorst E, Xiao S (2012) Sensitivity of the Chinese version of Hopkins Verbal Learning Test and Mini-Mental State Examination to Dementia and Demographics. The 15th asia-Pacific Regional Conference of Alzheimer's Disease.

37. HogervorstE, MursjidF, Irawati IR, PrasetyoS, MochtarN, NinukT, BandelowS, KusdhanyLS,, & Rahardjo TBW(2011) Validation of Two Short Dementia Screening Tests in Indonesia. Vascular Dementia: Risk Factors, Diagnosis and Treatment, 235-256.

38. Tanaka T (2005) Gender and ethnic differences on select verbal and visuospatial measures among older European and Japanese Americans. Pacific Graduate School of Psychology, 123: 3438.

39. Schwartz BS, Glass TA, Bolla KI, Stewart WF, Glass G, et al. (2004) Disparities in cognitive functioning by race/ethnicity in the Baltimore Memory Study. Environ Health Perspect 112: 314-320.

40. Prince M, Acosta D, Chiu H, Scazufca M, Varghese M; 10/66 Dementia Research Group (2003) Dementia diagnosis in developing countries: a cross-cultural validation study. Lancet 361: 909-917.

41. Tremont G, Miele A, Smith MM, Westervelt HJ (2010) Comparison of verbal memory impairment rates in mild cognitive impairment. J Clin Exp Neuropsychol 32: 630-636.

42. Norman MA, Evans JD, Miller WS, Heaton RK (2000) Demographically corrected norms for the California Verbal Learning Test. J Clin Exp Neuropsychol 22: 80-94.

43. Vanderploeg RD, Schinka JA, Jones T, Small BJ, Graves AB, et al. (2000) Elderly norms for the Hopkins Verbal Learning Test-Revised. Clin Neuropsychol 14: 318-324.