The Australian Psychological Society Ltd

Hopkins Verbal Learning Test: Normative data for older Australian adults

ROBERT L. HESTER[1,2] GLYNDA J. KINSELLA[1,3], BEN ONG[1], & MARGARET TURNER[1]

[1]*School of Psychological Science, La Trobe University, Melbourne, Victoria, Australia,* [2]*Department of Psychology, Trinity College, Dublin, Ireland, and* [3]*Department of Psychology, Caulfield General Medical Centre, Melbourne, Victoria, Australia*

Abstract
Normative data from 203 community-dwelling older Australian adults are presented in relation to performance on the Hopkins Verbal Learning Test. Age and years of education were identified as significantly impacting on performance. Therefore, data are presented for three separate age groups (60–69 years; 70–79 years; 80–89 years) and two different lengths of education groups. Comparison between the performance of the Australian sample and an available U.S. sample indicates a generally comparable profile of performance on the Hopkins Verbal Learning Test.

The Hopkins Verbal Learning Test (HVLT) (Brandt, 1991) and the revised version (HVLT-R) (Benedict, Schretlen, Groninger, & Brandt, 1998) are brief verbal memory and learning tests. The HVLT consists of a 12-item word list drawn from three semantic categories, presented in three learning trials that are followed by a *yes/no* recognition trial, containing the 12 target words and 12 distracter words (6 semantically related and 6 semantically unrelated). The HVLT-R also includes a 20–25 min delay followed by a delayed free-recall trial and a delayed recognition trial. The availability of six alternate forms, reduced number of list items and learning trials in comparison to other verbal learning measures, such as the Rey Auditory Verbal Learning Test, and sensitivity to dementia (Frank & Byrne, 2000), have been argued to make the HVLT particularly suitable for serial examinations and assessment of more severely impaired individuals (Lacritz, Cullum, Weiner, & Rosenberg, 2001). The HVLT is one of the most frequently used memory tests in clinical neuropsychological assessment of older adults (Vanderploeg et al., 2000).

Recent studies indicate that the HVLT has acceptable construct and concurrent validity (Shapiro, Benedict, Schretlen, & Brandt, 1999), and good interform and test–retest reliability (Benedict

et al., 1998). Normative data were recently published for a large community-dwelling American sample (*n* = 394) of older adult participants (Vanderploeg et al., 2000). Recent studies from the United States (Lacritz et al., 2001), United Kingdom (Hogervorst et al., 2001) and Australia (Frank & Byrne, 2000) have also recommended the HVLT as a valid and reliable screening test for mild dementia, indicating sensitivity levels ranging from .87 to .98 and specificity levels of .80 to .98, depending on the total score cut-off point used. For example, the Australian study (Frank & Byrne, 2000) demonstrated in a sample of patients with mild dementia that the HVLT detected 96% of cases using a cut-off HVLT total learning score of 18.

The aim of the present study was to provide normative data for the HVLT for a sample of older community-dwelling Australian adults, thereby extending the established clinical utility of the test in Australian contexts. Previous studies examining the performance of Australian samples on cognitive measures designed and normed in the United States have yielded some discrepant results (Casey & Heath, 1988; Shores & Carstairs, 2000; Worrall, Hickson, Barnett, & Yiu, 1998). For example, the Macquarie University Neuropsychological Normative Study (MUNNS) normative data provided by

Shores and Carstairs highlight the large discrepancy between U.S. and Australian normative data for the Wechsler Memory Scale, including a discrepancy of 10 points on the verbal memory index. These studies argue, on the basis of the discrepant results, that Australian norms are required for greater sensitivity of clinical interpretations. Given the reported sensitivity of the HVLT (and HVLT-R) to discriminate between older adults and patients with Alzheimer's disease, normative data for the HVLT were felt to be particularly useful for clinicians administering the HVLT to Australian populations.

Method

Participants

Data for the present study were collected as part of a series of experiments at La Trobe University School of Psychological Science investigating prospective remembering in older adults who were community-dwelling volunteers from metropolitan Melbourne. Participants were excluded on the basis of a premorbid history of neurological disorder, psychiatric illness, drug abuse or current use of psychoactive medication, or a Mini-Mental State Examination (MMSE) score < 25 (Folstein, Folstein, & McHugh, 1975). The sample comprised 203 participants, age ranging between 60 and 89 years ($M = 73.08$ years, $SD = 5.56$). The male : female ratio was 91:112, which was reasonably balanced given the increased life expectancy of Australian women over their male counterparts (Australian Bureau of Statistics, 2003). The mean years of education were 11.07 years ($SD = 3.10$ years), which is lower than that reported in previous U.S. studies providing normative data for older adults performance on the HVLT. For example, Vanderploeg et al. (2000) reported a sample mean of 14.1 mean years of education (SD was not provided), while Benedict et al. (1998) reported a sample mean of 13.8 years ($SD = 2.3$ years) of education.

Procedure

Participants were given a battery of neuropsychological measures that included the HVLT. Standard procedure for Form 1 of the HVLT was administered to all 203 participants. A subgroup of 136 participants also completed a delay procedure utilising the HVLT-R protocol. Following the three learning trials and the immediate recognition trial, a 20–25-min delay was interspersed before a delayed free-recall trial, followed by a delayed recognition trial as in the standard format for HVLT-R.

Mean scores for Trial 1, Trial 2, Trial 3, total score (sum of Trials 1–3), learning index (higher of Trial 2 or 3 minus Trial 1), and immediate discrimination index (true positives minus false positives) were calculated for all participants, while for the subgroup of 136 participants who received the HVLT-R delay procedure, mean scores for delayed recall, percent retention (best performance on Trial 2 or 3 minus delayed recall score, multiplied by 100) and delayed discrimination index were also calculated. Confidence intervals were calculated for each of the indices and presented for each age group Based on two standard errors either side of the current mean, they estimate the range within which 95% of all means of the same sample size would be predicted to fall (Chelune, 2002).

Results

The demographic data for the sample are presented in Table I. An initial treatment of the data investigated the influence of demographic variables: age, gender, and education on different measures of the HVLT-R.

A series of multiple regression analyses was used to examine whether demographic variables significantly influenced performance on total score, delayed recall, learning index, percent retention, and immediate and delayed discrimination index. Age and education were found to be significant predictors of total score, $F(3,199) = 7.61$, $MSE = 210.98$, $p < .01$, and delayed recall performance, $F(3,132) = 5.26$, $MSE = 49.54$, $p < .01$, and education was a predictor of the delayed discrimination index, $F(3,132) = 3.90$, $MSE = 17.40$, $p < .05$, although learning index, percent retention and the immediate discrimination indices were not significantly predicted by age, gender or education.

Given the significant effect of age on performance of the HVLT-R, the present sample's performance was divided into three age-groups: 60–69 years, 70–79 years and 80–89 years and these are presented in Tables 2–4. Performance of the age groups was also presented by years of education. These age group categories were selected to allow comparison to the previous published U.S. norms for older adults performance on the HVLT-R (Vanderploeg et al., 2000), as well as a subsample from this group matched for education to the Australian sample. The subsample included 167 participants with less than 13 years of education ($M = 11.3$, $SD = 3.0$) from the Vanderploeg et al. (2000) data, having a mean age of 72.4 years ($SD = 6.3$ years). A comparison of these samples is presented in Table V. A series of independent group t tests compared the mean performance of the Australian sample to the U.S. samples on each of the indices, indicating significant differences (at $p < .01$) for Trial 1, learning index, delayed recall

Table I. Demographic information for the Australian sample

| | | Demographic Variables | | | | | |
| | | Age | | Years of education | | Gender | |
Group	N	M	SD	M	SD	Male	Female
60–69	64	66.7	1.8	11.5	3.5	32	32
70–79	108	74.4	2.7	10.6	2.8	45	63
80–89	31	81.8	1.9	11.8	3.2	14	17
Total sample	203	73.1	5.6	11.1	3.1	91	112

Table II. Normative data for the HVLT-R performance of an Australian sample aged 60–69 years, further divided by education

| | | | | Education (years) | | | | | |
| | Total (n = 64) | | | ⩽ 10 (n = 29) | | | ⩾ 11 (n = 35) | | |
	M	SD	95%CI	M	SD	95%CI	M	SD	95%CI
Trial 1	5.8	1.7	5.4–6.3	5.2	1.5	4.6–5.8	6.4	1.7	5.8–6.9
Trial 2	7.7	2.3	7.1–8.3	6.7	2.2	6.0–7.5	8.5	2.0	7.8–9.1
Trial 3	9.0	2.1	8.5–9.6	8.1	2.3	7.3–8.9	9.8	1.6	9.1–10.5
Learning	3.3	1.6	2.9–3.7	2.9	1.8	2.3–3.5	3.5	1.5	3.0–4.1
Total Score	22.5	5.6	21.2–23.9	20.0	5.5	18.1–21.9	24.6	4.8	22.9–26.4
Immediate DI	10.4	2.0	9.8–10.9	10.0	2.6	9.2–10.8	10.8	0.9	9.9–11.6
	Total (n = 41)			⩽ 10 (n = 16)			⩾ 11 (n = 25)		
Delayed Recall	7.6	3.2	6.6 – 8.6	6.3	3.3	4.7–7.8	8.4	2.9	7.2–9.7
Percent Retained	79.3	28.3	70.4 – 88.2	73.7	34.8	58.0–89.3	82.9	23.3	70.3–95.5
Delay DI	9.3	2.1	8.6–10.0	8.4	2.3	7.3–9.5	9.9	1.8	9.0–10.8

Note. HVLT-R = Hopkins Verbal Learning Test – Revised; CI = confidence interval; DI = Discrimination Index.

Table III. Normative data for the HVLT-R performance of an Australian sample aged 70–79 years, further divided by education

| | | | | Education (years) | | | | | |
| | Total (n = 108) | | | ⩽ 10 (n = 63) | | | ⩾ 11 (n = 45) | | |
	M	SD	95%CI	M	SD	95%CI	M	SD	95%CI
Trial 1	5.0	1.8	4.6–5.3	4.7	1.9	4.3–5.1	5.3	1.5	4.8–5.8
Trial 2	6.8	1.9	6.4–7.1	6.8	2.0	6.3–7.3	6.7	1.9	6.1–7.3
Trial 3	8.0	2.3	7.6–8.4	7.8	2.5	7.3–8.4	8.2	2.0	7.6–8.9
Learning	3.2	1.8	2.9–3.5	3.2	1.8	2.8–3.7	3.1	1.8	2.6–3.6
Total Score	19.7	5.3	18.7–20.7	19.4	5.8	18.1–20.7	20.2	4.6	18.7–21.7
Immediate DI	10.5	2.0	10.1–11.0	10.5	1.6	10.0–11.1	10.5	2.4	9.8–11.1
	Total (n = 75)			⩽ 10 (n = 43)			⩾ 11 (n = 32)		
Delayed Recall	6.8	3.2	6.1–7.5	6.4	3.5	5.5–7.3	7.3	2.7	6.3–8.4
Percent Retained	81.7	31.4	74.5 – 89.0	80.4	36.3	70.8–90.0	83.6	23.9	72.5–94.7
Delay DI	9.1	2.1	8.6 – 9.6	8.9	2.2	8.3–9.6	9.4	2.0	8.6–10.2

Note. HVLT-R = Hopkins Verbal Learning Test – Revised; CI = confidence interval; DI = Discrimination Index.

and the delayed discrimination index between the Australian and total U.S. samples, but only Trial 1 remained significantly different when comparing the education matched groups.

Discussion

The results of the study were encouraging in support of the view that the performance of older Australian adults who were given the HVLT were similar to

Table IV. Normative data for the HVLT-R performance of an Australian sample aged 80–89 years, further divided by education

| | Total (n = 31) | | | Education (years) | | | | | |
| | | | | ≤ 10 (n = 15) | | | ≥ 11 (n = 16) | | |
	M	SD	95%CI	M	SD	95%CI	M	SD	95%CI
Trial 1	4.7	1.4	4.1–5.2	4.2	1.5	3.4–5.0	5.1	1.3	4.2–5.9
Trial 2	6.8	1.8	6.1–7.4	6.1	1.8	5.1–7.1	7.4	1.6	6.4–8.4
Trial 3	8.0	2.4	7.0–8.8	7.1	2.5	6.0–8.2	8.6	2.1	7.5–9.7
Learning	3.5	1.5	3.0–4.1	3.2	1.6	2.3–4.1	3.8	1.3	3.0–4.7
Total Score	19.5	5.0	17.4–21.1	17.4	5.2	14.8–20.0	21.1	4.6	18.5–23.6
Immediate DI	10.4	1.9	9.7–11.2	9.9	2.4	8.9–11.0	11.0	1.0	9.9–12
	Total (n = 20)			≤ 10 (n = 10)			≥ 11 (n = 10)		
Delayed Recalll	5.4	2.8	4.1–6.7	5.4	3.1	3.5–7.4	5.4	2.6	3.5–7.4
Percent Retained	68.1	38.8	50.0–86.3	78.3	48.8	58.5–98.2	57.9	23.6	38.0–77.8
Delay DI	8.8	2.6	7.5–10.1	9.0	2.3	7.6–10.4	8.6	3.1	7.1–10.1

Note. HVLT-R = Hopkins Verbal Learning Test–Revised; CI = confidence interval; DI = Discrimination Index

Table V. Comparison of education matched Australian (n = 203) and U.S. (n = 167) samples, and noneducation-matched Vanderploeg et al. (2000) U.S. sample (n = 394)

| | Australian sample (n = 203) | | U.S. sample (n = 394) | | U.S. education matched (n = 167) | | |
Variable	M	SD	M	SD	M	SD	p
Trial 1	5.2	1.7	4.8	1.7	4.7	1.8	**
Trial 2	7.1	2.1	7.4	2.1	7.2	2.4	
Trial 3	8.3	2.3	8.4	2.2	8.1	2.4	
Learning	3.3	1.7	3.8	1.8	3.7	1.8	*
Total Score	20.6	5.5	20.6	5.2	20.0	5.9	
Delayed Recall	6.9[a]	3.2	7.8	2.7	7.4	3.0	*
% Retained	79.0[a]	31.8	89.0	25.8	87.7	29.5	*
Delay DI	9.1[a]	2.2	9.7	1.9	9.6	2.1	*

Note. HVLT-R = Hopkins Verbal Learning Test–Revised; CI = confidence interval; DI = Discrimination Index.
[a]Australian sample for delay indices was n = 136.
*Significant difference between Australian and nonmatched U.S. sample (p < .01).
**Significant difference between Australian and both U.S. samples (p < .01).

their age-matched counterparts from the United States, particularly for the learning trials. It should be noted, however, that there is some variation across studies in the demographic variables found to impact upon levels of performance. Vanderploeg et al. (2000), in their sample of older U.S. adults, reported that age and gender impacted upon learning performance; and Benedict et al. (1998), in a general adult U.S. sample, reported that age and education impacted on learning performance. Within our Australian sample we found that age and education, but not gender, affected HVLT performance. The impact of age and education is observed in Tables 2–5 and it is seen that there is a clear decline in performance as age increases and years of education reduces. The lack of influence from gender on performance of the Australian sample was also surprising given the large impact it had in

the Vanderploeg et al. (2000) study, although such a finding is not without precedence. Two other large U.S. normative data studies (Benedict et al., 1998; Brandt, 1991) have also not reported gender effects on HVLT, and in the present study regression statistics suggested that gender effects were largely accounted for by age.

A direct comparison of the performances reported by Vanderploeg et al. (2000) and the present Australian sample revealed a generally lower level of performance on the delay indices. For example, our results demonstrate that older adults were retaining 10% less over a delay than their U.S. counterparts, which amounts to one fewer word remembered from an almost identical final learning trial performance. However, it is important to consider that the average education of the Vanderploeg et al. (2000) sample was 14.1 years, compared

to the present sample's 11.1 years, and that the regression analyses reported for our present sample indicated that education was a significant predictor of both total score and delayed recall performance. Subsequently, by constraining education to below 13 years, Vanderploeg et al. (2002, personal communication, 6 November) reported a comparable performance to our Australian sample for all but the Trial 1 performance. This result demonstrates how previously significant discrepancies between the samples can be ameliorated by adjusting for education differences.

We have provided both age- and education-adjusted normative data for the HVLT (and HVLT-R), given the performance differences associated with these factors. However, this also presents a limitation with interpreting some of the data, in particular the delayed recall indices and data for the oldest age group 80–89 years, for whom individual cell size is relatively low. A further constraint on the data is that all participants completed Form 1 of the HVLT, and although previous studies have demonstrated the equivalence of the six different forms (Benedict et al. 1998; Brandt, 1991), caution should be used when comparing performance on other forms to the present data.

Finally, future research with this test might provide additional data on age-related test–retest reliability to extend the normative data that now exists for the HVLT, and provide additional support to clinical decisions based on serial testing using the alternative forms of the test, particularly in the light of recent findings that indicate that the HVLT is sensitive to early cognitive change.

References

Australian Bureau of Statistics. (2003). 2001 Census Basic Community Profile and Snapshot. http://www.abs.gov.au (2003, November 22).

Benedict, R. H. B., Schretlen, D., Groninger, L., & Brandt, J. (1998). Hopkins Verbal Learning Test-Revised: Normative data and analysis of inter-form and test–retest reliability. *Clinical Neuropsychologist, 12*(1), 43–55.

Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *Clinical Neuropsychologist, 5*(2), 125–142.

Casey, P. J., & Heath, R. A. (1988). Category norms for Australians. *Australian Journal of Psychology, 40*(3), 323–339.

Chelune, G. (2002). Assessing reliable neuropsychological change. In R. D. Franklin (Ed.), *Prediction in forensic and neuropsychology* (pp. 123–147). Mahwah, NJ: Lawrence Erlbaum.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189–198.

Frank, R. M., & Byrne, G. J. (2000). The clinical utility of the Hopkins Verbal Learning Test as a screening test for mild dementia. *International Journal of Geriatric Psychiatry, 15*(4), 317–324.

Hogervorst, E., Combrinck, M., Lapuerta, P., Rue, J., Swales, K., & Budge, M. (2001). The Hopkins Verbal Learning Test and screening for dementia. *Dementia and Geriatric Cognitive Disorders, 13*(1), 13–20.

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

Shapiro, A. M., Benedict, R. H., Schretlen, D., & Brandt, J. (1999). Construct and concurrent validity of the Hopkins Verbal Learning Test–Revised. *Clinical Neuropsychologist, 13*(3), 348–358.

Shores, E. A., & Carstairs, J. R. (2000). The Macquarie University Neuropsychological Normative Study (MUNNS): Australian norms for the WAIS-R and WMS-R. *Australian Psychologist, 35*(1), 41–59.

Vanderploeg, R. D., Schinka, J. A., Jones, T., Small, B. J., Graves, A. B., & Mortimer, J. A. (2000). Elderly norms for the Hopkins Verbal Learning Test–Revised. *Clinical Neuropsychologist, 14*(3), 318–324.

Worrall, L., Hickson, L., Barnett, H., & Yiu, E. (1998). An evaluation of the Keep on Talking program for maintaining communication skills into old age. *Educational Gerontology, 24*(2), 129–140.