Research Article

Cross-cultural evaluation of learning and memory using a consonant-vowel-consonant trigram list

Prince Ampofo[1,2], Jessica L. Katschke[1], Kylie R. Kadey[1], Bradley J. Dixon[1], Colt M. Halter[1] ⬤, Allison C. Moll[1],
Maria Gattuso[2], Francesca Morganti[2] and John L. Woodard[1] ⬤

[1]Department of Psychology, Wayne State University, Detroit, MI, USA and [2]Department of Human and Social Sciences, University of Bergamo, Bergamo, Italy

Abstract

Objective: Word list-learning tasks are commonly used to evaluate auditory-verbal learning and memory. However, different frequencies of word usage, subtle meaning nuances, unique word phonology, and different preexisting associations among words make translation across languages difficult. We administered lists of consonant-vowel-consonant (CVC) nonword trigrams to independent American and Italian young adult samples. We evaluated whether an auditory list-learning task using CVC nonword trigrams instead of words could be applied cross-culturally to evaluate similar learning and associative memory processes. **Participants and Methods:** Seventy-five native English-speaking (USA) and 104 native Italian-speaking (Italy) university students were administered 15-item lists of CVC trigrams using the Rey Auditory Verbal Learning Test paradigm with five study-test trials, an interference trial, and short- and long-term delayed recall. Bayesian *t* tests and mixed-design ANOVAs contrasted the primary learning indexes across the two samples and biological sex. **Results:** Performance was comparable between nationalities on all primary memory indices except the interference trial (List B), where the Italian group recalled approximately one item more than the American sample. For both nationalities, recall increased across the five learning trials and declined significantly on the postinterference trial, demonstrating susceptibility to retroactive interference. No effects of sex, age, vocabulary, or depressive symptoms were observed. **Conclusions:** Using lists of unfamiliar nonword CVC trigrams, Italian and American younger adults showed a similar performance pattern across immediate and delayed recall trials. Whereas word list-learning performance is typically affected by cultural, demographic, mood, and cognitive factors, this trigram list-learning task does not show such effects, demonstrating its utility for cross-cultural memory assessment.

Keywords: nonword consonant-vowel-consonant trigrams; learning and memory; cross-cultural neuropsychology; interference

(Received 13 June 2023; final revision 6 September 2023; accepted 13 September 2023)

Over the last century, numerous memory assessment procedures have been developed to measure and differentiate normal from abnormal memory functioning. Most of these measures are highly face valid and focus on quantifying the amount of to-be-remembered information that can be recalled immediately after presentation and after a specified time delay. Most of these approaches are also highly verbal. Such procedures might evaluate the number of details from prose passages, the number of paired words, or the number of items on a list that can be recalled. Even recall of geometric figures can be susceptible to verbal labeling of shapes, suggesting that some linguistic processing occurs during purportedly visuospatial memory tasks. While most of these approaches effectively differentiate normal from abnormal memory functioning, performance on these measures can be influenced by demographic characteristics, such as sex, age, education, cultural and linguistic factors, and an examinee's familiarity with the to-be-recalled test material. For example,

females commonly show at least a slight verbal episodic memory performance advantage over males (Asperholm et al., 2019; Hirnstein et al., 2023). These variables can influence an individual's processing efficiency and may lead to memory scores contaminated by demographic, cultural, and linguistic differences.

Though basic cognitive mechanisms are considered similar cross-culturally (Nell, 1999), the behavioral manifestations of higher-order processes are undeniably influenced by an individual's culture (Fernández & Abe, 2018; Puente & Agranovich, 2004; Rivera Mindt et al., 2010). The influence of culture on cognitive performance typically favors individuals born and raised in the geographical region where a test was developed (Cole, 1998). Moreover, specific cognitive tasks may be more complex and require more cognitive and neural resources in one culture than another (Gutchess et al., 2011). In addition to these culture-specific effects, individual cognitive measures may assess different cognitive abilities depending on the examinee's cultural

background (Fasfous et al., 2013). Most cognitive tests are developed, standardized, and normed in predominantly "Western" and industrialized regions, such as the USA and UK, which share similar languages and origins and have high social and economic development rates. Several studies have reported that cross-cultural variability in socioeconomic and health status, as well as inequalities in educational opportunities, are confounds that can strongly influence cognitive test performance (Chin et al., 2012; Ferraro, 2016; Krch et al., 2015; Rosselli & Ardila, 2003; Schwartz et al., 2004; Weuve et al., 2018).

An essential cultural factor related to neuropsychological assessment is the language the examinee speaks. Because of the substantial verbal demands imposed by most memory measures, the words comprising a word list likely have shared and unique meanings to examinees who speak a specific language that do not lend themselves to translation. Additionally, word frequencies may differ substantially across languages. Simple word-for-word translations from one language to another may also change word phonology, such as the number of syllables per word, and subtle differences in semantic meanings of words, which collectively may influence memory performance. These limitations make direct translations of word list items from one language to another flawed and minimally informative.

Conventional word list memory measures usually use highly familiar words or material frequently occurring in the examinee's language. Familiarity with test stimuli could give some examinees an advantage on the task and may produce ceiling effects. Historically, Hermann Ebbinghaus (1885, 1964) made use of consonant-vowel-consonant (CVC) trigrams, a structured way of creating nonsense syllables, as stimuli in his quest to develop material that was devoid of inter-item associations and prior experience (Thorne & Henley, 2001). Nonword CVC trigrams are pronounceable combinations of letters (consonant-vowel-consonant) with no meaning or associations with other nonsense syllables. Nonword repetition depends more on the temporary storage of phonological representations in short-term memory during initial learning due to limited access to long-term lexical models that facilitate the recall of unfamiliar items. This manipulation eliminates the ability to use preexisting meaning, knowledge, or experience to facilitate recall.

To our knowledge, only one group has used CVC material within a list-learning task in two neuropsychological contexts. Bourke et al. (2012) presented the CVC nonsense syllables utilizing the structure of the Auditory Verbal Learning Test (AVLT; Rey, 1958) as part of a larger neuropsychological battery to study neuropsychological differences among persons with major depressive and social anxiety disorders, relative to matched controls. Persons with depression showed lower recall of CVC items across Trials 1–5 and a flatter learning curve than persons with social anxiety and matched controls. Delayed recall and recognition did not differ between groups. In a second study, Vierck et al. (2015) used the same paradigm to evaluate and screen persons for mild cognitive impairment. They reported that the CVC list-learning task showed similar psychometric characteristics as traditional list-learning tasks but with a reduced tendency for a ceiling effect. As in the first study, the CVC task was also sensitive to depression. These two studies suggest that using the AVLT paradigm with CVC material can evaluate processes associated with learning and memory, with concomitant reductions in the likelihood of ceiling effects. Importantly, this group does not appear to have investigated this procedure's potential for cross-cultural applications across speakers of different languages.

In this study, we implemented a similar CVC list-learning procedure to that of Bourke et al. (2012) and Vierck et al. (2015)

with independent samples of undergraduate students collected in the United States and Italy. Although the United States and Italy are considered "Western" and industrialized (Masuda et al., 2020), the predominant languages used in each country, a fundamental component of their respective cultures, differ phonologically and semantically. CVC trigrams eliminate the semantic aspect of words, and they hold the phonological structure constant across items. Beyond language, other critical cultural differences that could plausibly impact cognitive function exist between the US and Italy, such as access to education and health care (Petrelli et al., 2020), nutritional status (Zhang et al., 2015), climate, and other material, social, and cultural resources (Marks et al., 2006). Thus, we investigated whether similar learning and memory processes would be observed using CVC stimuli in native speakers of two different languages from North America and Italy. If similar learning material across languages can evaluate the same memory processes, such a finding could lead to developing a standardized cross-cultural memory task. We hypothesized that both groups of participants would show a comparable and increasing number of items recalled over the five learning trials, a similar number of items retrieved on the first learning trial and the interference trial, and an equivalent ability to retain CVC material over a 20-min delay. We also hypothesized that the American and Italian samples would perform equally on all performance indexes.

Method

Participants

The American sample included 75 (21 males, 54 females) Wayne State University undergraduate psychology students. The Italian sample consisted of 104 undergraduate psychology students (38 males, 66 females) from the University of Bergamo. Participants were recruited through the online Sona System at Wayne State University or an online Google form at the University of Bergamo. All participants received research credit for participation. Exclusion criteria included reporting any current or previous history of cognitive or neurological deficit, major psychiatric disorders, or any concern that would make participating in the research study challenging. Informed consent was obtained from all participants after fully explaining the research protocols and before starting the experimental session. The individual studies were approved by the Institutional Review Boards of Wayne State University and the University of Bergamo, and the research was completed in accordance with the Helsinki Declaration.

Measures

During the 20-min delay, both groups completed a demographic and health history questionnaire. Salthouse's Synonym and Antonym Test (Salthouse, 1993) and the Center for Epidemiologic Screening-Depression Scale (CES-D) (Radloff, 1977) were also administered during the delay to examine general intellectual functioning and depressive symptoms, respectively, for the American sample only. No parallel standardized Italian measures were available for the Italian participants, but they were given unstandardized direct translations of the Synonym and Antonym test and CES-D to fill the 20-min time delay. The Synonym and Antonym Test presents 20 target words with five response options for each word. Examinees must choose the correct synonym for the first ten target words, followed by selecting the correct antonym of the remaining ten target words. The CES-D consists of 20 statements reflecting depressive symptoms rated

Table 1. CVC trigrams for List A and List B used for the American and Italian studies

American		Italian	
List A	List B	List A	List B
HEV	VAB	HEV	VAB
SIJ[a]	REH	NIR[a]	REH
YIC[a]	GAQ	GOQ[a]	GAQ
FEH	XAW[b]	FEH	DAZ[b]
ZEW[a]	HIB	LEB[a]	HIB
ZUQ[a]	JOR[b]	SUQ[a]	ZAN[b]
KAJ[a]	KES[b]	BAV[a]	TEB[b]
JAL[a]	KIG[b]	ZAH[a]	VIH[b]
TOZ	WUF[b]	TOZ	QOL[b]
VUB[a]	TEV	VUD[a]	TEV
VAF	TOV	VAF	TOV
ZIM	VUM	ZIM	VUM
WUH[a]	GUZ	HUF[a]	GUZ
ZOJ[a]	ZIV	ZOG[a]	ZIV
GOX[a]	SOF	CIZ[a]	SOF

[a]Denotes List A trigrams that differed between nationality samples.
[b]Denotes List B trigrams that differed between nationality samples.

regarding their frequency of occurrence during the past week in four ordinal categories (Rarely/None of the Time to Most of the Time). We evaluated the possible contribution of these two constructs to performance on the CVC learning task, given that depression has reportedly been associated with performance on the CVC task (Bourke et al., 2012; Vierck et al., 2015) and crystallized intelligence is related to long-term memory (Hundal & Horn, 1977).

The CVC free recall task's structure used the AVLT framework. The AVLT is multi-trial list-learning approach for assessing immediate and delayed episodic memory, attention, and concentration (Lezak et al., 2012; Magalhães & Hamdan, 2010). Trigrams were presented using the software program, PsychoPy (Version V3.0.0b11; Peirce et al., 2019) on a Macintosh 27-inch 2015 iMac computer for the American study and an Apple MacBook Pro (13-inch, Mid 2012) for the Italian study. The learning material consisted of 15 test trigrams (List A) and 15 interference trigrams (List B). Each trigram consisted of three English letters in a consonant-vowel-consonant pattern. Candidate items were randomly generated and selected. Trigrams that were actual words, meaningful acronyms, popular abbreviations, homophones, or challenging to pronounce were excluded. The items were displayed in black, bold Arial font on a white background on the computer screen to enhance illumination.

Because there are only 21 letters in the Italian alphabet (the letters j, k, w, x, and y are not used), some of the original CVC trigrams that had originally been developed for the American group were substituted with new items for the Italian sample (10 items for List A and five for List B). Table 1 shows the items used for each list for the two groups.

Procedure

The American study involved one in-person study visit between May and December of 2019. The Italian study took place during the COVID-19 pandemic via an online Skype video call using screen sharing of the list words. For the American study, participants were seated directly opposite the examiner, with the computer monitor in a room illuminated by overhead fluorescent light. Before the start of the experiment, the experimenter briefed

the participant on the study's aims with the study information sheet. Informed consent was obtained, and all participants received a copy of this document. To begin the task, the participant faced the computer screen. Participants were seated about 91.44 cm (3 ft) from the computer screen. For the Italian online study, the experimenter briefed the participant on the study's aims and presented an informed consent form to the participant. Once participants agreed to participate in the study after reading the informed consent, the study began. Participants were asked to be in a quiet environment and indicated whether the screen was visible before the experiment commenced.

The CVC trigrams used for the American and Italian samples adopted the AVLT paradigm (Rey, 1958), replacing list words with CVC trigram nonwords. A 15-item list of nonword CVC trigrams (List A) was presented in a fixed order on five successive study-test recall trials. Trigrams appeared at a rate of one every 2 s, and participants read each item aloud as it was presented. At the end of each trial, participants were given a fixed period of 60 s to recall as many items as possible, and precisely 60 s separated the presentation of each trial. After the fifth trial, List B (consisting of a new 15-item list of nonword CVC trigrams) was presented, and participants were given 60 s to recall as many items as possible. Immediately after the recall of List B, participants were asked to identify as many trigrams from List A as possible within 60 s (Trial 6). Participants then completed a demographic questionnaire, the Salthouse Synonyms and Antonyms test, and the CES-D to fill a 20-min delay period. Next, for the delayed recall trial, participants were asked to recall as many items from list A as possible within 60 s. Performance indices included the number of items recalled on each trial (Trials 1 through 6, List B, and 20-min Delayed Recall), and the sum of items recalled across Trials 1 through 5.

Data analysis

Bayesian independent group *t* tests, correlations, and mixed-design analyses of variance (ANOVAs) were used to analyze the data. All analyses were conducted with JASP (JASP Team, 2023), version 0.17.1. A Bayesian analytic approach quantified the population parameter estimates most likely to underlie the observed data. The Bayes factor (BF_{10}) evaluates the degree of support for the alternate hypothesis relative to the null hypothesis given the observed data. BF_{10} values between 3 and 10 suggest anecdotal evidence favoring the alternate hypothesis, while values greater than 10 strongly support the alternate hypothesis. In contrast, increasing support for the null hypothesis is indicated as BF_{10} values become smaller than 1/3. As Bayesian effect size indexes for ANOVAs are difficult to compute using available software, frequentist partial η^2 effect sizes are presented to convey the magnitudes of the main and interaction effect sizes. The JASP default priors were as follows: ANOVAs used the uniform distribution as the prior; contingency table prior concentration was set to 1; *t* tests used the Cauchy prior scale 0.707; correlations used a stretched beta prior width of 1. Bayesian independent groups *t*-tests report the BF_{10}, the median Standardized Mean Difference (SMD) effect size, and the 95% Credible Interval (CI) of parameter estimates likely to have given rise to the data. Small, medium, and large SMD effect size benchmarks are typically considered to be 0.2, 0.5, and 0.8, respectively. Small, medium, and large partial η^2 effect size benchmarks are 0.01, 0.06, and 0.14, respectively. Because Bayesian analysis primarily focuses on the posterior distributions that are unaffected by the number of comparisons one wishes to make, and

Table 2. Analysis of effects of trial, nationality, and sex on number of words recalled on each trial for Bayesian mixed-design ANOVA

Effects	P(incl)	P(excl)	P(incl|data)	P(excl|data)	BF_{incl}
Trial	0.737	0.263	1.000	6.994×10^{-15}	$5.106 \times 10^{+13}$
Nationality	0.737	0.263	0.437	0.563	0.277
Sex	0.737	0.263	0.263	0.737	0.127
Trial × Nationality	0.316	0.684	0.003	0.997	0.006
Trial × Sex	0.316	0.684	0.002	0.998	0.005
Nationality × Sex	0.316	0.684	0.030	0.970	0.067
Trial × Nationality × Sex	0.053	0.947	6.907×10^{-8}	1.000	1.243×10^{-6}

Note: P(incl): prior probability associated with the plausibility of an effect before looking at the data; *P*(excl): prior probability associated with the implausibility of an effect before looking at the data; *P*(incl|data): posterior probability associated with the plausibility of an effect given the data; *P*(excl|data): posterior probability associated with the implausibility of an effect given the data; BF_{inclu}: the change from the prior inclusion odds to the posterior inclusion odds for each effect, averaged by all models including the effect, broadly reflecting the support for the effect given the data.

p-values are not used, corrections for multiple comparisons are unnecessary (Kruschke, 2015).

Results

Demographic characteristics

The Italian sample ($M_{age} = 21.3$ years, $SD = 2.7$ years, range = 18–29 years) was slightly older than the American sample ($M_{age} = 20.3$ years, $SD = 2.7$ years, range = 18–30 years; $BF_{10} = 3.7$, SMD effect size = 0.37, 95% Credible Interval (95% CI) = 0.1–0.7). The American sample mean CES-D score was 15.6 ($SD = 10.9$, range = 0–56, 95% CI = 13.1–18.1) and the mean Synonym-Antonym score was 6.4 out of 20 ($SD = 3.3$, range = 1–16, 95% CI = 5.6–7.2). A Bayesian test of independence between nationality (Italy, USA) and biological sex produced a BF_{10} of 0.392, suggesting approximately equal distributions of men and women across the two studies.

Correlations between learning indexes, age, CES-D, and synonym-antonym test

For the overall sample ($N = 179$), there was no evidence of a relationship between age and performance on any learning trial or learning index, as evidenced by extremely small correlation coefficients (all *r*'s < 0.1) and BF_{10}s (range = 0.09–0.24). Within the American sample only, none of the traditional learning measures correlated with the CES-D: Trial 1: $r = 0.12$ ($BF_{10} = 0.2$, 95% CI = −0.21 to 0.23); List B: $r = 0.02$ ($BF_{10} = 0.15$, 95% CI = −0.20 to 0.24); Trial 6: $r = −0.12$ ($BF_{10} = 0.24$, 95% CI = −0.41 to 0.02); Sum of Trials 1–5: $r = −0.17$ ($BF_{10} = 0.42$, 95% CI = −0.38 to 0.06); Delayed Recall: $r = −0.21$ ($BF_{10} = 0.72$, 95% CI = −0.41 to 0.02). Similarly, in the American sample, performance on the Synonym-Antonym Test was uncorrelated with these indexes: Trial 1: $r = 0.11$ ($BF_{10} = 0.22$, 95% CI = −0.12 to 0.32); List B (interference): $r = 0.03$ ($BF_{10} = 0.15$, 95% CI = −0.20 to 0.25); Trial 6: $r = 0.23$ ($BF_{10} = 0.96$, 95% CI = 0.0 to 0.43); Sum of Trials 1–5: $r = 0.26$ ($BF_{10} = 1.72$, 95% CI = 0.03–0.45); Delayed Recall: $r = 0.21$ ($BF_{10} = 0.74$, 95% CI = −0.02 to 0.41). Given the absence of correlations between the learning measures and age, depressive symptoms, or crystallized intelligence, they were not considered further as potential covariates.

Learning over trials

A 2 (Nationality: Italy, USA) × 2 (Sex: Male, Female) × 5 (Learning Trial: Trial 1 through Trial 5) Bayesian mixed-design ANOVA examined learning over trials. The dependent variable was the number of items correctly recalled on each trial. Of 19 possible models, including each combination of main and/or interaction effects, a model containing only the Learning Trial main effect had the highest BF_{10} ($BF_{Model} = 13.22$, partial $\eta^2 = .56$). This finding indicates that there was no evidence for main effects of Sex or Nationality, nor were there any other interaction effects, as illustrated in the model-averaged results table for each main and interaction effect in Table 2. Partial η^2 values for all other main and interaction effects were 0.01 or less. The BF_{10} for the model including the Learning Trial main effect was $5.1 \times 10^{+13}$, which indicates the robust evidence in the data for the model that only includes Learning Trial. The BF_{10}s for the other main and interaction effects, were less than 0.28. Notably, there were no appreciable effects of Sex or Nationality on the learning curve.

Comparison of American and Italian samples on memory performance indexes

Table 3 presents the means, SDs, BF_{10}s, and standardized mean group difference effect sizes from a series of independent groups Bayesian *t*-tests for each of the five learning trials, List B, Trial 6, Delayed Recall, and the sum of words recalled across Trials 1–5 for the American and Italian participants, separately. Figure 1 presents the means and 95% credible intervals for each index by each nationality. Except for performance on List B, no group differences on any other performance index were observed, as indicated by the minimal BF_{10} values. Italian participants performed slightly better than the American participants on List B by approximately one item. Except for List B, which showed a medium effect size, all performance indexes showed conventionally small or near-zero effect sizes.

Comparison of male and female performance on memory performance indexes

Table 4 presents the descriptive statistics for the AVLT learning and memory indexes separately for men and women, collapsed over nationality. Figure 2 displays the means and 95% credible intervals for each index in graphical form. As can be gleaned from the table, the small BF_{10} values suggest no group differences in memory performance between men and women. All effect sizes are near zero or in the conventionally small range.

Susceptibility to proactive and retroactive interference

Susceptibility to proactive interference was investigated using a 2 (Nationality: Italy, USA) × 2 (Trial: Trial 1, List B) mixed-design Bayesian ANOVA. The model-averaged results in Table 5 demonstrate a robust main effect of Nationality (inclusion

Table 3. Means and standard deviations for learning indexes by nationality

| | Group | N | Mean | SD | SE | 95% credible interval | | BF$_{10}$ | Median Effect Size | 95% credible interval for effect size | |
						Lower	Upper			Lower	Upper
Trial 1	Italy	104	4.3	1.7	0.2	3.9	4.6	0.96	0.28	−0.01	0.57
	USA	75	3.8	1.4	0.2	3.5	4.1				
Trial 2	Italy	104	6.3	2.4	0.2	5.8	6.7	0.50	0.22	−0.07	0.51
	USA	75	5.7	2.3	0.3	5.2	6.3				
Trial 3	Italy	104	7.6	2.4	0.2	7.1	8.1	0.64	0.24	−0.05	0.53
	USA	75	7.0	2.3	0.3	6.5	7.5				
Trial 4	Italy	104	8.5	2.7	0.3	8.0	9.1	0.22	0.11	−0.17	0.40
	USA	75	8.2	2.6	0.3	7.6	8.8				
Trial 5	Italy	104	9.5	2.5	0.2	9.0	10.0	0.37	0.19	−0.10	0.48
	USA	75	9.0	2.6	0.3	8.4	9.6				
List B	Italy	104	4.5	1.8	0.2	4.1	4.8	51.2	0.51	0.21	0.81
	USA	75	3.5	1.6	0.2	3.2	3.9				
Trial 6	Italy	104	6.8	2.7	0.3	6.3	7.4	0.21	−0.10	−0.38	0.19
	USA	75	7.1	3.0	0.3	6.5	7.8				
DR	Italy	104	7.2	2.8	0.3	6.6	7.7	0.22	−0.11	−0.40	0.18
	USA	75	7.5	2.9	0.3	6.9	8.2				
T1–T5 sum	Italy	104	36.3	10.0	1.0	34.3	38.2	0.64	0.24	−0.05	0.53
	USA	75	33.7	9.7	1.1	31.5	35.9				

Note: DR = Delayed Recall; T1–T5 Sum = Sum of words recalled across Trials 1–5; *SD* = standard deviation; *SE* = Standard Error; 95% Credible Interval reflects the lower and upper bounds of the 95% most likely mean value for each nationality; 95% Credible Interval for Effect Size reflects the upper and lower bounds for the 95% most likely standardized mean difference effect sizes.
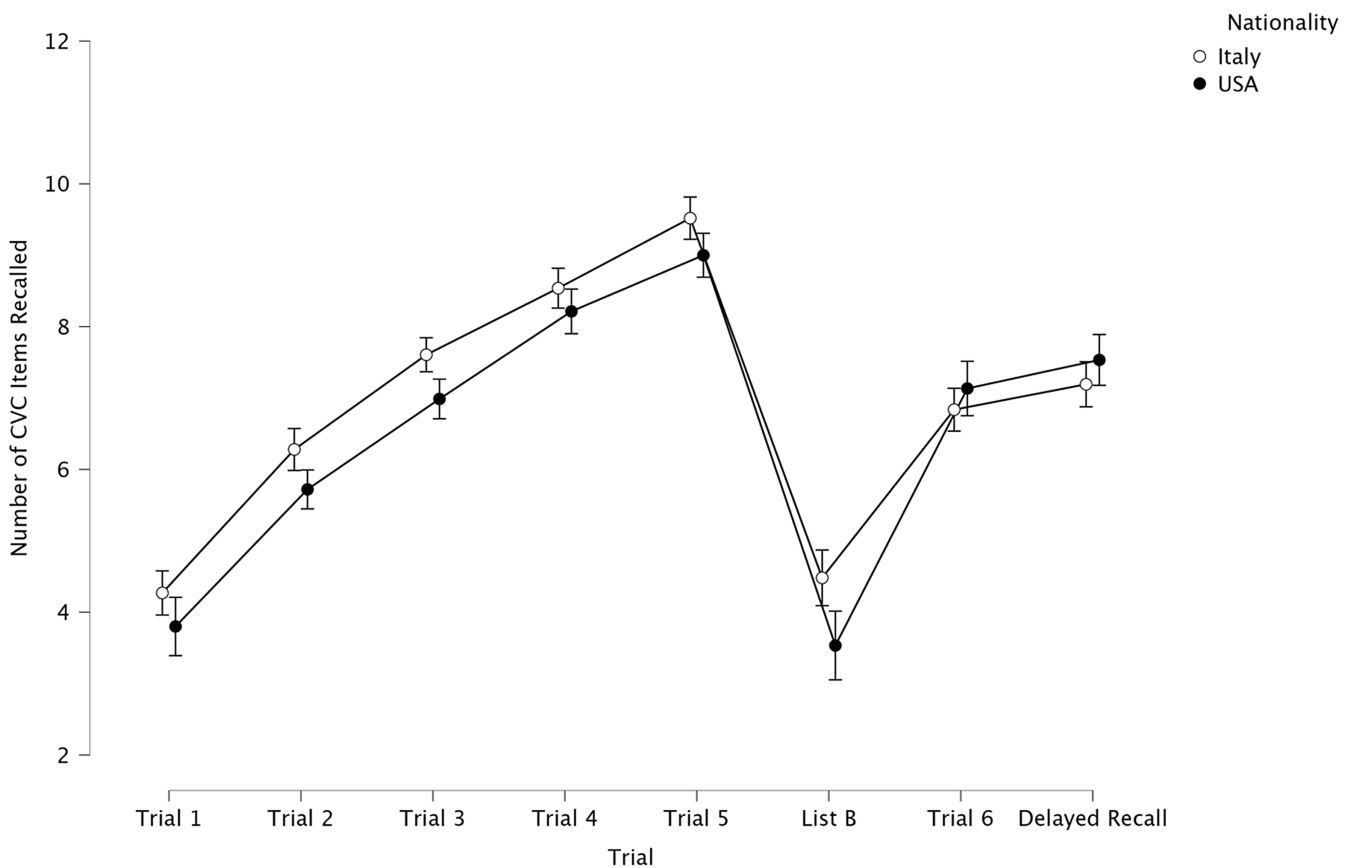
Figure 1. Means and 95% credible intervals for learning and memory indexes by nationality.

BF$_{10}$ = 20.8, partial $\eta^2 = 0.06$), whereby the Italian sample performed better than the American sample, averaged over Trial 1 and List B. There was no evidence of susceptibility to proactive interference for the overall sample (Trial main effect inclusion BF$_{10}$ = 0.1, partial $\eta^2 = 2.1 \times 10^{-4}$) or a different pattern for the two Nationalities (Nationality × Trial interaction effect inclusion BF$_{10}$ = 0.3, partial $\eta^2 = 0.02$). These results are presented in Figure 3, whereby the Nationality main effect seems to be driven by the previously observed difference between the Italian and American samples, primarily on List B.

Table 4. Means and standard deviations for learning indexes by sex

	Group	N	Mean	SD	SE	95% credible interval Lower	95% credible interval Upper	BF$_{10}$	Median Effect Size	95% credible interval for effect size Lower	95% credible interval for effect size Upper
Trial 1	Female	120	4.0	1.5	0.1	3.7	4.3	0.20	−0.08	−0.38	0.22
	Male	59	4.2	1.7	0.2	3.7	4.6				
Trial 2	Female	120	6.0	2.4	0.2	5.5	6.4	0.21	−0.10	−0.40	0.20
	Male	59	6.2	2.4	0.3	5.6	6.9				
Trial 3	Female	120	7.3	2.3	0.2	6.8	7.7	0.22	−0.10	−0.40	0.20
	Male	59	7.5	2.5	0.3	6.9	8.2				
Trial 4	Female	120	8.2	2.5	0.2	7.8	8.7	0.34	−0.18	−0.48	0.12
	Male	59	8.7	3.0	0.4	8.0	9.5				
Trial 5	Female	120	9.2	2.7	0.2	8.8	9.7	0.19	−0.07	−0.37	0.23
	Male	59	9.4	2.4	0.3	8.8	10.1				
List B	Female	120	3.9	1.8	0.2	3.6	4.2	0.63	−0.25	−0.55	0.05
	Male	59	4.4	1.9	0.2	3.9	4.9				
Trial 6	Female	120	6.8	2.8	0.3	6.3	7.4	0.23	−0.12	−0.42	0.18
	Male	59	7.2	2.8	0.4	6.5	7.9				
DR	Female	120	7.3	2.7	0.2	6.8	7.8	0.17	−0.03	−0.33	0.27
	Male	59	7.4	3.1	0.4	6.6	8.2				
T1–T5 sum	Female	120	34.8	9.7	0.9	33.0	36.5	0.24	−0.13	−0.43	0.17
	Male	59	36.1	10.5	1.4	33.4	38.9				

Note: DR = Delayed Recall; T1–T5 Sum = Sum of words recalled across Trials 1–5; SD = standard deviation; SE = Standard Error; 95% Credible Interval reflects the lower and upper bounds of the 95% most likely mean value for each nationality; 95% Credible Interval for Effect Size reflects the upper and lower bounds for the 95% most likely standardized mean difference effect sizes.
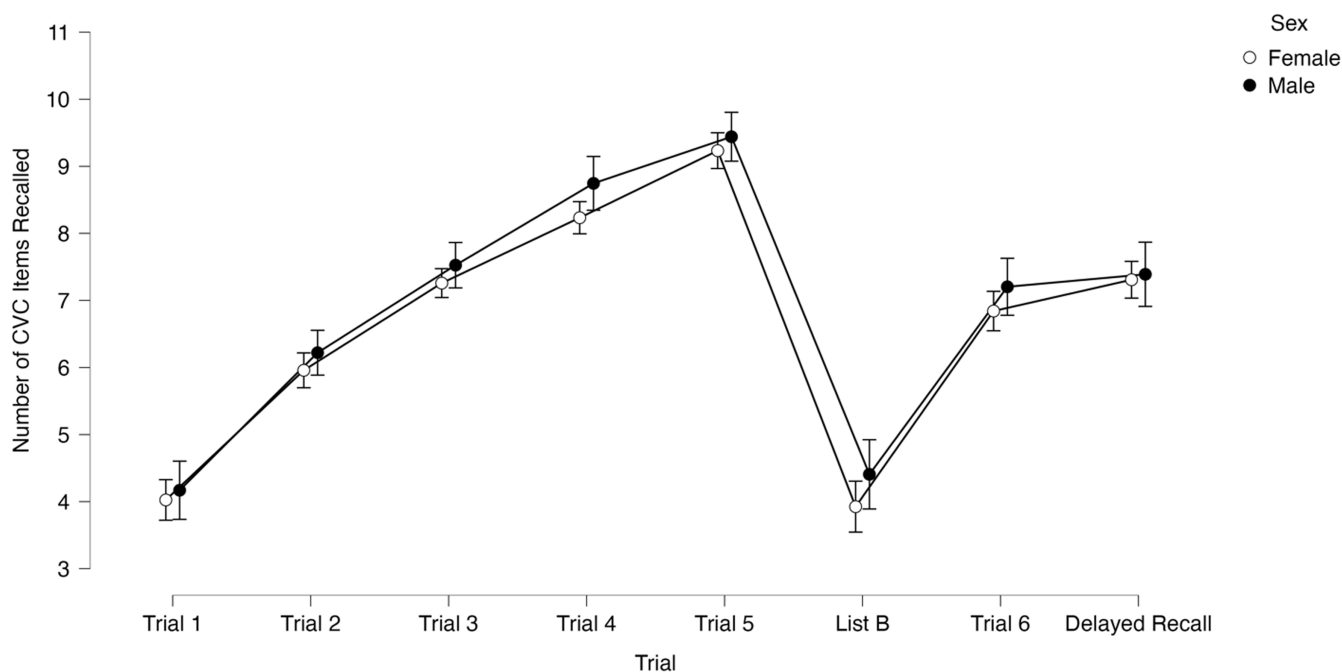
Figure 2. Means and 95% credible intervals for learning and memory indexes by sex.

Another 2 (Nationality: Italy, USA) × 2 (Trial: Trial 5, Trial 6) Bayesian ANOVA examined possible pre and postinterference differences in recall that would reflect susceptibility to retroactive interference. Table 6 displays the model-averaged effects suggesting a robust main effect of Trial (inclusion BF$_{10}$ = 1.4 × 10^{+14}, partial η^2 = 0.56), no main effect of Nationality (inclusion BF$_{10}$ = 1.0; partial η^2 = 4.8 × 10^{-4}), and anecdotal evidence suggesting an interaction effect (inclusion BF$_{10}$ = 3.9, partial η^2 = 0.04). Figure 4 illustrates the means and 95% credible intervals for the two Nationality groups on the preinterference (Trial 5) and postinterference (Trial 6) recall trials. As indicated by the overlapping 95% credible intervals for the two Nationality groups, the evidence favoring an interaction does not appear compelling. In contrast, there does appear to be robust evidence in favor of susceptibility to retroactive interference for both groups, given the substantial performance decline following the presentation of the interference list.

Discussion

The primary aim of this cross-cultural study was to determine whether a list of nonword CVC trigrams presented in the

Table 5. Analysis of effects of trial, nationality, and trial × nationality for susceptibility to proactive interference (List A Trial 1 vs. List B Trial)

Analysis of effects					
Effects	P(incl)	P(excl)	P(incl|data)	P(excl|data)	BF_incl
Trial	0.600	0.400	0.156	0.844	0.123
Nationality	0.600	0.400	0.969	0.031	20.843
Trial × Nationality	0.200	0.800	0.061	0.939	0.261

Note: P(incl): prior probability associated with the plausibility of an effect before looking at the data; P(excl): prior probability associated with the implausibility of an effect before looking at the data; P(incl\|data): posterior probability associated with the plausibility of an effect given the data; P(excl\|data): posterior probability associated with the implausibility of an effect given the data; BF_incl: the change from the prior inclusion odds to the posterior inclusion odds for each effect, averaged by all models including the effect, broadly reflecting the support for the effect given the data.
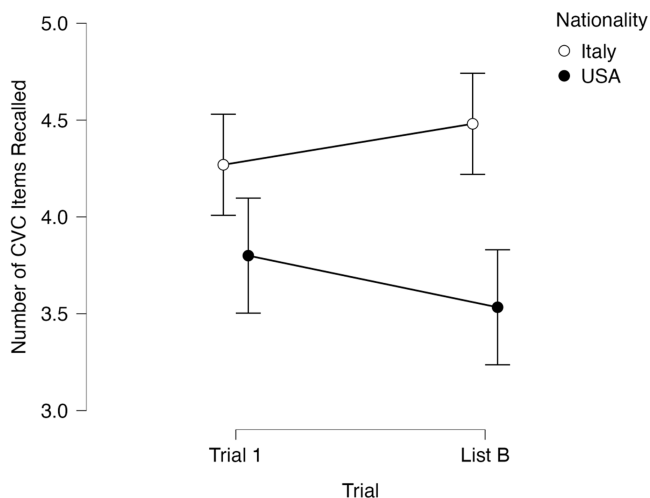
Figure 3. Means and 95% credible intervals for nationality × trial Bayesian ANOVA for susceptibility to proactive interference.

framework of the AVLT (Rey, 1958) could be used to evaluate learning and memory for native speakers of different languages. Our results suggested that this approach can assess auditory learning and delayed recall in native English and Italian speakers. No performance differences were observed between the Italian and American participants on any of the five learning trials, the postinterference recall trial, the sum of Trials 1–5, or the Delayed Recall trial. The Italian group performed better than the American participants only on List B by approximately one item. No evidence for susceptibility to proactive interference was observed for either group. In contrast, robust susceptibility to retroactive interference was observed for both groups. Susceptibility to retroactive interference is not commonly observed on the AVLT for healthy individuals in this age group, but it is more commonly seen after age 60 (Vakil et al., 2010). No performance differences were observed between men and women, which was surprising given the typical verbal memory advantage often observed for women (Asperholm et al., 2019; Crossley et al., 1997; Hirnstein et al., 2023). In addition, in the American sample, none of the memory outcome measures were related to age, depressive symptoms, as indexed by the CES-D, or crystallized intelligence, as measured by the Salthouse Synonym and Antonym Test.

Language is the most prominent issue when attempting to establish the cross-cultural equivalence of a standardized assessment. Standardized learning and memory assessments have historically tended to be heavily language-based and were developed primarily in "Western" cultures, thus biasing assessment results (Mushquash & Bova, 2007). For example, the most commonly used learning and memory assessments are auditory verbal list-learning tasks, such as the AVLT paradigm (Rey, 1958). Ideally, simple translations of the assessment items and instructions would yield comparable psychometric properties. However, previous research suggests that several significant considerations must be made during translation, and psychometric equivalence is not ensured (Cromer et al., 2013; Rendu et al., 2012). Simple translations, especially for heavily language-based tasks or complex verbal instructions, can be problematic because word meaning and usage vary as a function of language and culture (Ardila, 2021). As a result, the direct translation of word learning tasks seems to be insufficient due to factors such as preexisting semantic connections, frequency and familiarity, different cultural approaches to testing and conceptualizations of memory, and other cultural nuances (Ardila, 2021; Leger & Gutchess, 2021). These factors become further complicated when considering the role of language functioning in bilingual individuals (Rivera Mindt et al., 2008). Even within "Western" nations, ethnic diversity has dramatically increased in recent decades as international immigration has become more accessible. As a result, neuropsychologists must be prepared to encounter and assess individuals from diverse backgrounds (Franzen et al., 2021; Goudsmit et al., 2017). Thus, there is a pressing need for a more thorough development of standardized assessments of learning and memory that remain psychometrically useful across languages and, thus, cultures.
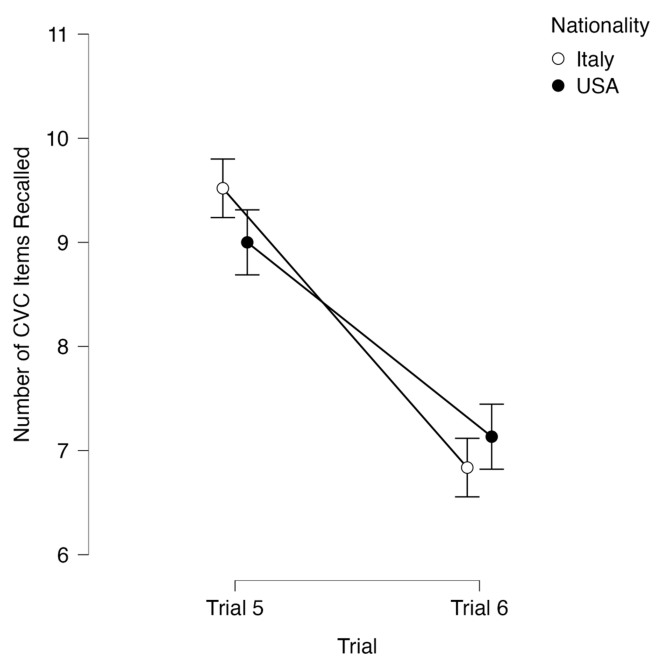
Our study offers a novel demonstration that performance on a verbal list-learning task of nonword CVC items was functionally equivalent across samples from the USA and Italy. Using nonword CVC items delivered in the AVLT paradigm precludes many issues inherent in language-based tasks, such as individual and cultural differences in preexisting semantic associations and familiarity with items. Thus, using unfamiliar nonwords standardized in form and length (e.g., CVC) reduces language-related factors, such as differences in phonology and semantics across languages, that may bias participant performance. Unlike most word list-learning tasks, participants read each item aloud as it was visually presented in our task. This procedure may have introduced a unique supportive learning process not seen in other measures. However, we still did not observe a ceiling effect. Despite these advantages, relatively little work has been done to modify the AVLT paradigm with nonword CVC items (Bourke et al., 2012; Vierck et al., 2015). The two previous studies using a similar approach to our current work were conducted with samples from New Zealand. With the addition of our research to the limited extant literature, there is evidence from three distinct cultures that using nonword CVC items in the AVLT paradigm is a valid measure of learning and memory, with participants performing similarly across cultures. The current results suggest that this modified paradigm is a suitable cross-cultural measure of auditory-verbal learning and memory, given that similar stimuli evoked roughly equivalent responses across two distinct cultures and languages.

No significant disparities between males and females or learning, interference, or delayed recall trials were observed. This finding was somewhat surprising, as females tend to perform better on verbal memory tasks than males (Crossley et al., 1997; Geffen et al., 1990; Gordon & Clark, 1974; Graves et al., 2017; Kimura & Seal, 2003; Kramer et al., 1988; Norman et al., 2000; Weiss et al., 2006; Woodard, 2006). Bleecker et al. (1988) found that women outperformed men on most AVLT performance indexes. Interestingly, Kimura and Seal (2003) found that females outperformed men in recalling actual

Table 6. Analysis of effects of trial, nationality, and trial × nationality for susceptibility to retroactive interference (List A Trial 5 vs. List A postinterference Trial 6)

	Analysis of effects				
Effects	P(incl)	P(excl)	P(incl|data)	P(excl|data)	BF$_{incl}$
Trial	0.600	0.400	1.000	4.885×10^{-15}	$1.365 \times 10^{+14}$
Nationality	0.600	0.400	0.604	0.396	1.017
Trial × Nationality	0.200	0.800	0.495	0.505	3.923

Note: P(incl): prior probability associated with the plausibility of an effect before looking at the data; P(excl): prior probability associated with the implausibility of an effect before looking at the data; P(incl\|data): posterior probability associated with the plausibility of an effect given the data; P(excl\|data): posterior probability associated with the implausibility of an effect given the data; BF$_{inclu}$: the change from the prior inclusion odds to the posterior inclusion odds for each effect, averaged by all models including the effect, broadly reflecting the support for the effect given the data.

Figure 4. Means and 95% credible intervals for nationality × trial Bayesian ANOVA for susceptibility to retroactive interference.

words but not nonsense words. One possible reason for the absence of sex differences in performance may have been due to the nonsemantic and associative nature of the stimulus items. Because the CVC trigrams are equally unfamiliar to men and women, neither group had an advantage, resulting in no sex disparity in recall performance. These results imply that female superiority in verbal memory may result from using items that are familiar to the examinee, have preexisting associations, or have a unique semantic salience (e.g., "bird" may be more memorable than "kestrel" due to its greater frequency of use and because "bird" represents a superordinate category).

Performance on the CVC trigram memory test was also unrelated to age. This finding is unsurprising due to the restricted age range consisting only of younger adults. Future studies using this approach should consider the possible effects of age in a sample that spans a more extensive age range. More surprising, however, was the absence of a relationship between crystallized intelligence and memory performance. As noted earlier, crystallized intelligence has a demonstrated association with long-term memory (Hundal & Horn, 1977). In a more recent study, Rapport et al. (1997) showed strong relationships between Verbal IQ on the Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981) and learning indexes on the Wechsler Memory Scale-Revised

(Wechsler, 1987) and California Verbal Learning Test (Delis et al., 1987) in a young adult sample. The strong relationships between crystallized intelligence and conventional verbal memory scores imply a performance advantage for individuals with larger vocabularies or preexisting familiarity with the to-be-learned material. The absence of associations between crystallized intelligence and all learning indexes on the CVC trigram memory test in the American sample suggests that performance on this task may be a purer index of auditory memory, as the CVC trigrams are devoid of semantic information that could potentially confer any performance advantage.

Depression has been inconsistently related to verbal memory performance. Some studies show relatively intact memory functioning in depression (Egeland et al., 2003; Hammar et al., 2011; Hammar & Årdal, 2013), and others report impaired memory functioning (Chen et al., 2018; Lee et al., 2012; Wang et al., 2022), particularly for individuals with recurrent depression (Basso & Bornstein, 1999). We found no relationships between CES-D score and memory performance on the CVC trigram test in the American sample, suggesting that it may be relatively insensitive to depressive symptoms. However, this finding should be interpreted cautiously, as our sample was relatively young and not a clinical group, which would likely restrict the range of CES-D scores. In addition, the two previous studies using the CVC trigram memory test (Bourke et al., 2012; Vierck et al., 2015) did observe modest but significant relationships with other measures of depression in older clinical and community samples (Bourke et al., 2012: mean age approximately 38 years, range 18–65 years; Vierck et al., 2015: mean age not reported, range 49–51 years) and they included participants meeting criteria for major depressive disorder. Thus, greater depression severity would be expected in those studies compared to our study where major depressive disorder was an exclusion criterion. The other two studies also used a faster presentation time than in our study (1 s/item versus 2 s/item), and they provided external auditory presentation of list items instead of having participants read the items aloud. Future cross-cultural research with this measure should evaluate the possible effects of depression on performance in both community and clinical samples with a broad age range.

The single significant nationality difference across individual trials reflected better List B performance for the Italian participants than for the American group. Although the mean difference amounted to only one word, it was associated with a medium effect size. While this isolated difference is puzzling, possible explanations could include a possible advantage conferred by online administration of the task or slightly better performance associated with the trigram set composition of List B for the Italian sample than the trigrams used for List B in the American group. A larger working memory capacity in the Italian sample relative to the American sample might be a third possibility that could be directly tested in future research.

Limitations and future directions

Though the results of this study provide preliminary evidence supporting the cross-cultural use of CVC trigrams to assess aspects of auditory-verbal learning and memory, several limitations must be considered. First, the administration contexts differed for the two nationality groups, as the American participants were tested in person. In contrast, the Italian participants were tested online due to the COVID-19 pandemic. However, there were no group differences on any learning index, apart from List B, and this difference was relatively modest. This finding seems to imply that the administration format did not systematically influence CVC trigram learning task performance. Second, the CVC word lists shared many to most of the same items, but they were not identical because the original list of CVC trigrams used with the American sample included trigrams containing certain letters not used in the Italian alphabet. Again, except for List B, no systematic performance differences were observed between the two groups despite having somewhat different trigrams, suggesting that the specific composition of the individual trigrams may have minimal to no influence on performance. Nevertheless, future work with the cross-cultural application of the CVC trigram learning task should consider differences in alphabet composition and other phonological differences across languages when constructing a list of trigrams. In doing so, a common list of items could be developed for use across several languages, assuming they use a common set of Latin/Roman alphabetical characters. This approach may be more challenging to apply in languages that use non-Latin/Roman alphabetical characters. Ease of pronunciation of CVC trigrams across languages should also be carefully considered. Some trigrams may be challenging to enunciate for speakers of some languages or dialects. Finally, the trigram memory task was not as limited by a ceiling effect as some word list-learning measures. However, CVC trigrams assess pure phonological verbal memory due to the nonsemantic nature of the items. The extent to which phonological verbal memory is sensitive to the effects of aging, neurodegenerative conditions, or brain injury as word list learning would be worthwhile to investigate in future research. A direct comparison of this novel task with existing word list-learning measures in the same participants would be helpful to determine similarities and differences between the learning processes tapped by each measure. Because the task removes semantic information but holds item phonology constant, it will be interesting to contrast this measure's sensitivity to preclinical Alzheimer's disease with existing word list-learning tasks. Finally, using this procedure with individuals from "non-Western" and less industrialized geographical regions would be an essential next step to establishing further the clinical utility of this approach.

Conclusion

Typical word list-learning measures that require examinees to recall words after the presentation of the word list commonly use semantically associated or nonassociated words or material already familiar to the learner. Because the word list stimuli typically include everyday items or high-frequency words, they may be less difficult to recall than nonsense syllables, especially among younger adults at the peak of their cognitive abilities. Nonword CVC trigrams are nonsemantic, eliminating familiarity with and preexisting associations among items. CVC trigram lists have a particular advantage over word lists in that the same common core set of trigrams can potentially be constructed to assess phonological memory across speakers of many different languages that use the same Latin/Roman alphabetical characters. The present study's findings suggest that using nonword CVC trigrams in a list-learning paradigm can overcome the challenges posed by preexisting familiarity and associations among words used in traditional list-learning measures of memory. This approach can also be used for cross-cultural memory assessment, as we found minimal differences in performance between Italian and American young adults. Moreover, the task appeared to assess learning and memory equally well for males and females, regardless of nationality. Finally, the trigram memory test reduces the likelihood of ceiling effects, minimizes sex differences in verbal memory performance and performance advantages conferred by extensive vocabulary knowledge, and lends itself to creating many different alternate forms.

References

Ardila, A. (2021). Cross-cultural differences in cognition and learning. In T. K. Shackleford (Ed.), *The SAGE handbook of evolutionary psychology: Foundations of evolutionary psychology* (pp. 420–435). Sage Reference.

Asperholm, M., Högman, N., Rafi, J., & Herlitz, A. (2019). What did you do yesterday? A meta-analysis of sex differences in episodic memory. *Psychological Bulletin*, 145(8), 785–821. https://doi.org/10.1037/bul0000197

Basso, M. R., & Bornstein, R. A. (1999). Relative memory deficits in recurrent versus first-episode major depression on a word-list memory task. *Neuropsychology*, 13(4), 557–563.

Bleecker, M. L., Bolla-Wilson, K., Agnew, J., & Meyers, D. A. (1988). Age-related sex differences in verbal memory. *Journal of Clinical Psychology*, 44(3), 403–411.

Bourke, C., Porter, R. J., Carter, J. D., McIntosh, V. V., Jordan, J., Bell, C., Frances, C., Colhoun, H. C., & Joyce, P. R. (2012). Comparison of neuropsychological functioning and emotional processing in major depression and social anxiety disorder subjects, and matched healthy controls. *Australian & New Zealand Journal of Psychiatry*, 46(10), 972–981.

Chen, C., Jiang, W. H., Wang, W., Ma, X. C., Li, Y., Wu, J., Hashimoto, K., & Gao, C. G. (2018). Impaired visual, working, and verbal memory in first-episode, drug-naive patients with major depressive disorder in a Chinese population. *PLoS One*, 13(4), e0196023.

Chin, A. L., Negash, S., Xie, S., Arnold, S. E., & Hamilton, R. (2012). Quality, and not just quantity, of education accounts for differences in psychometric performance between African Americans and white non-Hispanics with Alzheimer's disease. *Journal of the International Neuropsychological Society*, 18(2), 277–285.

Cole, M. (1998). *Cultural psychology: A once and future discipline.* Harvard University Press.

Cromer, J., Krishna, V., Nguyen, A., Acquadro, C., & Fuller, D. S. (2013). Cross-cultural adaptation of a research version of the Rey Auditory Verbal Learning Test (RAVLT) into Japanese. *Value in Health*, 16(3), A42.

Crossley, M., D'Arcy, C., & Rawson, N. S. (1997). Letter and category fluency in community-dwelling Canadian seniors: A comparison of normal participants to those with dementia of the Alzheimer or vascular type. *Journal of Clinical and Experimental Neuropsychology*, 19(1), 52–62.

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

Ebbinghaus, H. (1885). *Uber das Gediichtnis. Untersuchungen zur experimentellen Psychologie.* Duncker & Humblot.

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology.* Dover.

Egeland, J., Sundet, K., Rund, B. R., Asbjørnsen, A., Hugdahl, K., Landrø, N. I., Lund, A., Roness, A., & Stordal, K. I. (2003). Sensitivity and specificity of

memory dysfunction in schizophrenia: A comparison with major depression. *Journal of Clinical and Experimental Neuropsychology*, 25(1), 79–93. https://doi.org/10.1076/jcen.25.1.79.13630

Fasfous, A. F., Hildalgo-Ruzzante, N., Vilar-López, R., Catena-Martínez, A., & Pérez-García, M. (2013). Cultural differences in neuropsychological abilities required to perform intelligence tasks. *Archives of Clinical Neuropsychology*, 28(8), 784–790.

Fernández, A. L., & Abe, J. (2018). Bias in cross-cultural neuropsychological testing: Problems and possible solutions. *Culture and Brain*, 6(1), 1–35. https://doi.org/10.1007/s40167-017-0050-2

Ferraro, F. R. (2016). *Minority and cross-cultural aspects of neuropsychological assessment: Enduring and emerging trends* (2nd ed.). Taylor and Francis.

Franzen, S., Papma, J. M., van den Berg, E., & Nielsen, T. R. (2021). Cross-cultural neuropsychological assessment in the European Union: A Delphi expert study. *Archives of Clinical Neuropsychology*, 36(5), 815–830.

Geffen, G., Moar, K. J., O'Hanlon, A. P., Clark, C. R., & Geffen, L. B. (1990). Performance measures of 16- to 86-year-old males and females on the Auditory Verbal Learning Test. *The Clinical Neuropsychologist*, 4(1), 45–63.

Gordon, S. K., & Clark, W. C. (1974). Application of signal detection theory to prose recall and recognition in elderly and young adults. *Journal of Gerontology*, 29(1), 64–72.

Goudsmit, M., Uysal-Bozkir, Ö., Parlevliet, J. L., van Campen, J. P., de Rooij, S. E., & Schmand, B. (2017). The Cross-Cultural Dementia Screening (CCD): A new neuropsychological screening instrument for dementia in elderly immigrants. *Journal of Clinical and Experimental Neuropsychology*, 39(2), 163–172.

Graves, L. V., Moreno, C. C., Seewald, M., Holden, H. M., Van Etten, E. J., Uttarwar, V., McDonald, C. R., Delano-Wood, L., Bondi, M. W., Woods, S. P., Delis, D. C., & Gilbert, P. E. (2017). Effects of age and gender on recall and recognition discriminability. *Archives of Clinical Neuropsychology*, 32(8), 972–979. https://doi.org/10.1093/arclin/acx024

Gutchess, A. H., Schwartz, A. J., & Boduroğlu, A. (2011). The influence of culture on memory. In: *FAC 2011: Directing the future of adaptive systems – 6th international conference*, Orlando, FL, USA, 9–14 July 2011.

Hammar, Å., & Årdal, G. (2013). Verbal memory functioning in recurrent depression during partial remission and remission - Brief report. *Frontiers in Psychology*, 4. https://doi.org/10.3389/fpsyg.2013.00652

Hammar, Å., Isaksen, L., Schmid, M., Årdal, G., & Strand, M. (2011). MDD patients show intact memory performance when given optimal conditions. *Applied Neuropsychology*, 18(3), 191–196.

Hirnstein, M., Stuebs, J., Moè, A., & Hausmann, M. (2023). Sex/gender differences in verbal fluency and verbal-episodic memory: A meta-analysis. *Perspectives on Psychological Science*, 18(1), 67–90. https://doi.org/10.1177/17456916221082116

Hundal, P. S., & Horn, J. L. (1977). On the relationships between short-term learning and fluid and crystallized intelligence. *Applied Psychological Measurement*, 1(1), 11–21.

JASP Team (2023). *JASP* (Version 0.17.1) [Computer software].

Kimura, D., & Seal, N. B. (2003). Sex differences in recall of real or nonsense words. *Psychological Reports*, 93(1), 263–264.

Kramer, J., Delis, D., & Daniel, M. (1988). Sex differences in verbal learning. *Journal of Clinical Psychology*, 44, 907–915.

Krch, D., Lequerica, A., Arango-Lasprilla, J. C., Rogers, H. L., DeLuca, J., & Chiaravalloti, N. D. (2015). The multidimensional influence of acculturation on digit symbol-coding and Wisconsin card sorting test in Hispanics. *The Clinical Neuropsychologist*, 29(5), 624–638.

Kruschke, J. K. (2015). *Doing Bayesian data analysis: A tutorial with R, JAGS, and STAN* (2nd ed.). Academic Press.

Lee, R. S., Hermens, D. F., Porter, M. A., & Redoblado-Hodge, M. A. (2012). A meta-analysis of cognitive deficits in first-episode Major Depressive Disorder. *Journal of Affective Disorders*, 140(2), 113–124.

Leger, K. R., & Gutchess, A. (2021). Cross-cultural differences in memory specificity: Investigation of candidate mechanisms. *Journal of Applied Research in Memory and Cognition*, 10(1), 33–43.

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). Oxford University Press.

Marks, G. N., Cresswell, J., & Ainley, J. (2006). Explaining socioeconomic inequalities in student achievement: The role of home and school factors. *Educational Research and Education*, 12(2), 105–128.

Masuda, T., Batdorj, B., & Senzaki, S. (2020). Culture and attention: Future directions to expand research beyond the geographical regions of WEIRD cultures. *Frontiers in Psychology*, 11, Article 1394. https://doi.org/10.3389/fpsyg.2020.01394

Magalhães, S., & Hamdan, A. (2010). The Rey Auditory Verbal Learning Test: Normative data for the Brazilian population and analysis of the influence of demographic variables. *Psychology and Neuroscience*, 31, 85–91.

Mushquash, C. J., & Bova, D. L. (2007). Cross-cultural assessment and measurement issues. *Journal on Developmental*, 13(1), 53–65.

Nell, V. (1999). Luria in Uzbekistan: The vicissitudes of cross-cultural neuropsychology. *Neuropsychology Review*, 9(1), 45–52.

Norman, M. A., Evans, J. D., Miller, W. S., & Heaton, R. K. (2000). Demographically corrected norms for the California Verbal Learning Test. *Clinical and Experimental Neuropsychology*, 22(1), 80–94.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Petrelli, A., Di Napoli, A., Demuru, E., Ventura, M., Gnavi, R., Di Minco, L., Tamburini, C., Mirisola, C., & Sebastiani, G. (2020). Socioeconomic and citizenship inequalities in hospitalisation of the adult population in Italy. *PLoS ONE*, 15(4), e0231564. https://doi.org/10.1371/journal.pone.0231564

Puente, A. E., & Agranovich, A. V. (2004). The cultural in cross-cultural neuropsychology. In M. Hersen, G. Goldstein, & S. R. Beers (Eds.), *Comprehensive handbook of psychological assessment, Vol 1: Intellectual and neuropsychological assessment* (pp. 321–332). John Wiley & Sons, Inc.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1, 385–401.

Rapport, L. J., Axelrod, B. N., Theisen, M. E., Brines, D. B., Kalechstein, A. D., & Ricker, J. H. (1997). Relationship of IQ to verbal learning and memory: Test and retest. *Journal of Clinical and Experimental Neuropsychology*, 19(5), 655–666. https://doi.org/10.1080/01688639708403751

Rendu, E., Caveney, A., Miner, A., Nomikos, A., & Acquadro, C. (2012). Cross-cultural adaptation of a research version of the Rey Auditory Verbal Learning Test (RAVLT) into US Spanish. *Value in Health*, 15(4), A97–A98.

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

Rivera Mindt, M., Arentoft, A., Kubo Germano, K., D'Aquila, E., Scheiner, D., Pizzirusso, M., Sandoval, T. C., & Gollan, T. H. (2008). Neuropsychological, cognitive, and theoretical considerations for evaluation of bilingual individuals. *Neuropsychology Review*, 18, 255–268.

Rivera Mindt, M., Byrd, D., Saez, P., & Manly, J. (2010). Increasing culturally competent neuropsychological services for ethnic minority populations: A call to action. *Clinical Neuropsychologist*, 24(3), 429–453.

Rosselli, M., & Ardila, A. (2003). The impact of culture and education on non-verbal neuropsychological measurements: A critical review. *Brain and Cognition*, 52(3), 326–333. https://doi.org/10.1016/s0278-2626(03)00170-2

Salthouse, T. A. (1993). Speed and knowledge as determinants of adult age differences in verbal tasks. *Journal of Gerontology: Psychological Sciences*, 48(1), 29–36.

Schwartz, B. S., Glass, T. A., Bolla, K. I., Stewart, W. F., Glass, G., Rasmussen, M., Bressler, J., Shi, W., & Bandeen-Roche, K. (2004). Disparities in cognitive functioning by race/ethnicity in the Baltimore Memory Study. *Environmental Health Perspectives*, 112(1), 314–320. https://doi.org/10.1080/00224545.1978.9924096

Thorne, B. M., & Henley, T. B. (2001). *Connections in the history and systems of psychology* (2nd ed.). Houghton, Mifflin, and Company.

Vakil, E., Greenstein, Y., & Blachstein, H. (2010). Normative data for composite scores for children and adults derived from the Rey Auditory Verbal Learning Test. *The Clinical Neuropsychologist*, 24(4), 662–677.

Vierck, E., Porter, R. J., Spittlehouse, J. K., & Joyce, P. R. (2015). Evaluation of the consonant vowel consonant (CVC) test as an assessment and screening instrument for mild cognitive impairment. *International Journal of Psychological Studies*, 7(4), 119–129.

Wang, M., Xiongwei, Q., Yang, X., Fan, H., Dou, Y., Guo, W., Wang, Q., Chen, E., Li, T., & Ma, X. (2022). The pattern glare and visual memory are disrupted in patients with major depressive disorder. *BMC Psychiatry*, *22*(1), 518. https://doi.org/10.1186/s12888-022-04167-9

Wechsler, D. (1981). *Wechsler Adult lntelligence Scale-Revised*. The Psychological Corporation.

Wechsler, D. (1987). *Wechsler Memory Scale-Revised*. The Psychological Corporation.

Weiss, E. M., Ragland, J. D., Brensinger, C. M., Bilker, W. B., Deisenhammer, E. A., & Delazer, M. (2006). Sex differences in clustering and switching in verbal fluency tasks. *Journal of the International Neuropsychological Society*, *12*(04), 502–509. https://doi.org/10.1017/s1355617706060656

Weuve, J., Rajan, K. B., Barnes, L. L., Wilson, R. S., & Evans, D. A. (2018). Secular trends in cognitive performance in older black and white U.S. adults, 1993-2012: Findings from the Chicago Health and Aging Project. *Journal of Gerontology Series B, Psychological Sciences and Social Sciences*, *73*, S73–S81. https://doi.org/10.1093/geronb/gbx167

Woodard, J. L. (2006). Memory performance indexes for the Rey Auditory Verbal Learning test. In A. M. Poreh (Ed.), *The quantified process approach to neurological assessment* (pp. 105–141). Taylor & Francis.

Zhang, R., Wang, Z., Fei, Y., Zhou, B., Zheng, S., Wang, L., Huang, L., Jiang, S., Liu, Z., Jiang, J., & Yu, Y. (2015). The difference in nutrient intakes between Chinese and Mediterranean, Japanese, and American diets. *Nutrients*, *7*(6), 466–4688. https://doi.org/10.3390/nu7064661