Studies in Second Language Acquisition (2024), **46**, 581–596
doi:10.1017/S0272263124000093

METHODS FORUM

Do verbal and nonverbal declarative memory tasks in second language research measure the same abilities?

Phillip Hamrick, Christopher A. Was and Yin Zhang

Kent State University
Corresponding author: Phillip Hamrick; Email: phamric1@kent.edu.

(Received 07 February 2023; Revised 21 November 2023; Accepted 28 November 2023)

Abstract

A growing body of evidence demonstrates that individual differences in declarative memory may be an important predictor of second language (L2) abilities. However, the evidence comes from studies using different declarative memory tasks that vary in their reliance on verbal abilities and task demands, which preclude estimating the size of the relationship between declarative memory and L2 learning. To address these concerns, we examined the relationship between verbal and nonverbal declarative memory abilities within the same task while controlling for task demands and stimulus modality, to estimate the upper bound of the relationship between verbal and nonverbal declarative memory. Results indicate that when task demands and stimulus modality are controlled, verbal and nonverbal declarative memory abilities shared a medium-to-large amount of underlying variance. However, future studies should exercise caution in appraising associations between declarative memory abilities and L2 learning until a more precise understanding of the underlying mechanisms is achieved.

Introduction

A growing body of work on individual differences in second language (L2) learning has emphasized the importance of examining domain-general cognitive abilities (i.e., those that subserve multiple cognitive domains, not just language), especially those abilities that have demonstrable links to the neurocognitive systems that support learning and memory. Much of this work has focused on the roles of individual differences in declarative and procedural memory abilities (for overviews, see Hamrick et al., 2018; Morgan-Short et al., 2022). The present study focuses on the former, as several theories predict important roles for declarative memory in L2 learning and processing, as a system that plays a critical role (i) in the L2 lexicon (e.g., Paradis, 2009; Ullman, 2004, 2016, 2020; Witzel & Forster, 2012), (ii) in early stages of L2 learning across multiple linguistic domains (DeKeyser, 2020; Ullman, 2016, 2020), and (iii) in developing the kinds of metalinguistic knowledge that are a common part of classroom L2 learning (Paradis, 2009; Ullman, 2016).

Because of the wide range of roles attributed to declarative memory in L2, numerous studies have examined to what degree individual differences in declarative memory abilities are associated with L2 abilities. Across those studies, a diverse array of declarative memory tasks has been used. Some studies have employed tasks that operationalize declarative memory as recognition memory (e.g., Hamrick et al., 2019; Murphy et al., 2021; Zhang et al., 2021), while others have used tasks that operationalize declarative memory as associative learning (e.g., Artieda & Muñoz, 2016; Bowles et al., 2016; Buffington & Morgan-Short, 2018; Ettlinger et al., 2014; Granena, 2019; Hamrick, 2015; Li, 2017; Linck et al., 2013; Saito, 2017, 2019). Moreover, some studies have incorporated both as either individual predictors (Ruiz et al., 2018; Ruiz et al., 2021; Walker et al., 2020) or as composite scores (Faretta-Stutenberg & Morgan-Short, 2018; Morgan-Short et al., 2014; Pili-Moss et al., 2020). Across these studies, it has commonly been found that declarative memory abilities are predictive of several aspects of L2 learning, including vocabulary, grammatical, and phonological development (see Morgan-Short et al., 2022 for an overview).

In addition to varying in their task demands (recognition memory vs associative learning), the stimuli employed within these tasks vary in their modality (auditory/visual) and stimulus domain (verbal/nonverbal). Such unsystematic variability in task demands, modality, and stimulus domains across studies makes it difficult to reconcile discrepancies in the literature over the role of declarative memory in L2 learning and processing (Morgan-Short et al., 2022). Indeed, the task demands associated with recognition memory and associative learning appear to tap different neuroanatomical aspects of the declarative memory system (Eichenbaum, 2012). Moreover, variability in task demands, modality, and stimulus domain makes it difficult to ascertain reliable effect size estimates of the magnitude and direction of the relationship between individual differences in declarative memory and language abilities.

In the following sections, we underscore the importance of task demands and stimulus domain in measuring declarative memory abilities, raising particular concerns about the verbal nature of some declarative memory tasks. We then describe a study designed to provide an upper bound estimate of the association between verbal and nonverbal declarative memory abilities with hopes that it will lead to the improvement of the validity of declarative memory tasks in L2 research and, therefore, also lead to higher validity of individual differences research.

Declarative memory task demands

The declarative memory system is not monolithic and homogeneous in its structure or function (Eichenbaum, 2012). Rather, it subserves multiple interrelated cognitive abilities, two of which are often used as measures of declarative memory: recognition memory and associative learning. Recognition memory refers to the ability to recognize a previously encountered stimulus (usually experienced during an encoding phase of a study), based on either recollection (i.e., successful retrieval of the previous stimulus study episode) or a feeling of familiarity (Yonelinas, 2002). Rather than simply recognizing previously encountered items, associative learning tasks are used to measure how people encode and retrieve associations between stimuli (Arndt, 2012), such as word pairs or word-object pairs or stimuli and their context (Eichenbaum, 2012), such as words and other context-specific details like font color. Within L2 research, declarative memory measures based on associative learning tasks most typically employ word pairs (e.g., Modern Language Aptitude Test section V (MLAT-V): Carroll & Sapon, 1959) or word-object pairs (e.g., LLAMA-B: Meara, 2005; Applied Linguistics

Progr Ammes for the Computerised Assessment of Aptitude (ALPACCA): Rogers et al., 2023).

Critically, recognition memory and associative learning tasks appear to tap different underlying neural substrates within the declarative memory system. Associative learning and memory rely especially on the hippocampus, while item recognition appears to rely especially on the rhinal cortex when driven by familiarity but also the hippocampus when driven by recollection (Davachi, 2006; Eichenbaum, 2012; Yonelinas et al., 2007). Moreover, the perirhinal cortex and lateral entorhinal cortex play important roles in processing objects, people, and events, whereas the medial entorhinal cortex appears to encode information about spatial and temporal context. Thus, these substrates play independent, albeit interactive, roles in declarative memory, and different declarative memory tasks may differentially tap these substrates (e.g., simple item recognition may primarily tap the perirhinal and lateral entorhinal cortex, while registration of context while recruiting the medial entorhinal cortex, and processing of the associations between an item and its context recruits the hippocampus).

Inasmuch as declarative memory plays a role in the L2, these item recognition, context sensitive, and associative learning components of declarative memory may play different roles across different aspects of the L2 (e.g., in word recognition vs. word learning). It may, therefore, be unsurprising that studies that use recognition memory or associative learning tasks sometimes come to somewhat different conclusions about the role of declarative memory in L2 (see Morgan-Short et al., 2022 for an overview). Indeed, although recognition memory tasks and associative learning tasks should correlate with one another due to their interactions within the declarative memory system, the precise nature of the tasks (e.g., their differential reliance on recognition, familiarity, and context for processing stimuli) may increase or reduce the magnitude of the correlation.

Verbalness of individual difference measures

Adding to the confusion, commonly used recognition memory and associative learning tasks also vary in their use of verbal and nonverbal declarative memory tasks, creating a potentially serious confound in the findings. As Morgan-Short et al. (2022) argue, in studies of language, measures of individual cognitive differences should be nonverbal to avoid "shared verbalness." Specifically, while it is possible that (non)significant correlations between language abilities and a verbal declarative memory task could reflect a role (or not) for declarative memory mechanisms in L2 learning and processing, it is also entirely possible that correlations between the two could be due to the verbal nature of both the language and the declarative memory stimuli. The use of such verbal declarative memory tasks in L2 research is common, with numerous studies using verbal associative learning tasks as measures of declarative memory abilities (Artieda & Muñoz, 2016; Bowles et al., 2016; Buffington & Morgan-Short, 2018; Ettlinger et al., 2014; Granena, 2019; Hamrick, 2015; Li, 2017; Linck et al., 2013; Saito, 2017, 2019). This potential problem is not unique to L2 research, either. Indeed, research on declarative memory in psychology, cognitive science, neuroscience, and neuropsychology all routinely use verbal measures of associative learning, list memorization, or recognition/recall, despite the lack of clear evidence that verbal and nonverbal declarative memory rely on precisely the same underlying mechanisms.

Moreover, task demands and verbalness may be complicating factors when they are combined in certain types of declarative memory tasks, such as verbal paired-associate learning tasks. These tasks evaluate declarative memory abilities based on associative learning of word pairs or word-meaning pairs (e.g., MLAT-V, LLAMA-B), but they can be thought of, essentially, as word learning tasks. These types of tasks ask learners to memorize new words and either another word, a translation equivalent, or an object, mirroring the word learning processes inherent in language acquisition. Use of such tasks may make any statistically significant correlation between declarative memory and L2 abilities potentially unsurprising (e.g., word learning in a verbal paired-associate learning task is correlated with some aspect of L2 learning) or, worse, trivial (e.g., word learning in a verbal paired-associate learning task is correlated with L2 word learning). In either case, the shared verbal nature of the stimuli in such declarative memory tasks and in any L2 task creates a potential confound, limiting our understanding of the true role of declarative memory in the L2 as well as potentially overestimating the effect size of the relationship.

The present study

When some researchers assume that verbal declarative memory tasks tap the same underlying system as nonverbal declarative memory tasks, they are making an assumption about what is fundamentally an empirical, and hitherto unanswered, question: to what degree do verbal and nonverbal declarative memory abilities tap the same underlying abilities? The present study sought to begin addressing this question. Specifically, our aim was to establish the upper bound estimate of the degree to which verbal and nonverbal declarative memory tasks tap the same underlying processes. Bearing in mind that recognition memory and associative learning tasks[1] may tap partly dissociable declarative memory capacities, we sought to control for task demands by creating a declarative memory task that examines only one type of declarative memory (recognition memory) in both verbal and nonverbal domains.

In this initial study, we sought to answer the following question: What are the relationships between verbal and nonverbal declarative memory abilities (in this case, encoding and recognition memory abilities indexed by accuracy and reaction times) after controlling for task demands and stimulus modality? Given the theoretical predictions that memory for words should rely on the same broad declarative system that underlies memory for nonlinguistic materials (Ullman, 2016), we predicted that correlations between verbal and nonverbal declarative abilities would be significantly and positively correlated; however, we did not have explicit predictions about the magnitude of such correlations. Large correlations would indicate that verbal and nonverbal declarative memory abilities share mostly the same underlying mechanisms, while small or nonsignificant correlations would indicate that verbal and nonverbal declarative memory abilities rely on more distinct mechanisms.

Method

Participants

A priori power analysis based on the effect size estimate reported in Hamrick et al. (2018) of $r = .40$ indicated that for 90% power and an alpha = .05, 47 participants would be needed. Anticipating attrition or lost data, we recruited a total of fifty-four volunteer undergraduate students ($M_{age} = 20.63$, $SD_{age} = 1.68$, $\text{range}_{age} = 19\text{–}24$; three left-handed and one ambidextrous) at Kent State University who participated in the study in exchange for extra credit. Forty-five participants identified as female, the rest male.

---

[1] Some previous studies have reported correlations between the two measures ranging from $r = .149$ to $r = .38$: Buffington et al., 2021; Morgan-Short et al., 2014; Buffington & Morgan-Short, 2018).

All participants except one reported English as their native language; the other participant reported being English-Indian bilingual from birth.

Materials

Verbal and nonverbal materials

In the verbal-nonverbal declarative memory task, participants were visually presented with verbal stimuli (words and pseudowords in black font on a white background), which were taken from the English Lexicon Project website (Balota et al., 2007), and nonverbal stimuli (black line drawings of real and made-up objects on a white background), which were taken from the normed stimuli reported in Snodgrass and Vanderwart (1980) and Kroll and Potter (1984). There were 128 total items (32 words, 32 pseudowords, 32 drawings of real objects, 32 drawings of made-up objects) in the task. Sixty-four of these items (16 words, 16 pseudowords, 16 real objects, 16 made-up objects) were presented in the encoding phase, and all items were presented in the recognition phase. A list of the stimulus labels is available in the Supplemental Materials.

The recognition phase of the task (see description below) required that participants make old/new decisions to all 128 stimuli. To ensure that the old and new stimuli were not significantly different from one another in important psycholinguistic characteristics, we examined item-level values of the stimuli. For the 32 real words across the old/new distinction, no significant differences were observed between the number of phonemes ($F(1, 30) = .025, p = .876$), word length ($F(1, 30) = .018, p = .895$), body-object interaction ($F(1, 30) = 2.759, p = .107$), word frequency ($F(1, 30) = .388, p = .538$), or the number of orthographical neighbors ($F(1, 30) = 1.049, p = .314$). For the 32 pseudowords across the old/new distinction, no significant differences were observed between word length ($F(1, 30) = .092, p = .764$) or the number of orthographical neighbors ($F(1, 30) = 2.169, p = .151$). The descriptive statistics for verbal stimuli are presented in Table 1. To account for any effects of subvocalization on the encoding or recognition of the nonverbal picture stimuli, the word labels that corresponded to the pictures of the thirty-two real objects were also examined, with no significant differences observed between the number of phonemes ($F(1, 30) = .000, p = 1.000$), familiarity ($F(1, 30) = 2.875, p = .100$), body-object interaction ($F(1, 30) = .006, p = .939$), frequency ($F(1, 30) = 1.275, p = .268$), or nameability ($F(1, 30) = .788, p = .382$). Finally, the only norm that would impact subvocalization for the made-up objects was object nameability, but there

Table 1. Mean descriptive statistics for verbal stimuli in the verbal-nonverbal declarative memory task

Status	Number of phonemes	Word length	Word frequency	Number of orthographic neighbors	Body-object interaction rating
Real Old	4.6	5.8	2.4	3.9	3.3
Real New	4.6	5.7	2.4	2.3	3.9
Made-up Old	NA	5.9	NA	4.2	NA
Made-up New	NA	5.8	NA	2.6	NA

Note. Word frequencies were log-transformed values from the SUBTLEXus corpus as reported in the English Lexicon Project. NA = not available.

Table 2. Descriptive statistics for nonverbal stimuli in the verbal-nonverbal declarative memory task

Status	Number of phonemes	Familiarity rating	Body-object interaction rating	Word frequency	Nameability rating
Real Old (M)	4.8	4.8	5.0	2.4	4.9
Real New (M)	4.8	4.9	5.0	2.7	4.9
Made-up Old (M)	NA	NA	NA	NA	1.4
Made-up New (M)	NA	NA	NA	NA	1.5

Note. Word frequencies were log-transformed values from the SUBTLEXus corpus as reported in the English Lexicon Project.

was no significant difference between old and new made-up objects in nameability ($F(1, 30) = .585$, $p = .450$). The descriptive statistics for the nonverbal stimuli are presented in Table 2.

Verbal-Nonverbal Declarative Memory Task

The goal of the verbal-nonverbal declarative memory task was to assess the relationship between declarative memory abilities for verbal and nonverbal stimuli, while controlling for task demands and modality. To control for task demands, all stimuli were presented within the same task. To control for modality, all stimuli were presented visually as black objects (black text or black line drawings) on a white background. The task consisted of two main phases (an incidental encoding phase and a recognition phase) with a distractor task in between. The incidental encoding phase was assumed to tap semantic memory abilities (lexical and object decisions for known words and objects/concepts), while the recognition phase was assumed to tap episodic memory abilities (i.e., memory for what was recently presented), although semantic memory abilities could also play a role in recognition as well.

The goal of the incidental encoding phase was to require participants to encode stimuli (so that their memory for the stimuli could be subsequently tested in the recognition phase) while also eliciting a measure of their encoding ability (through making a category decision). During this incidental encoding phase, participants classified visually presented English words, pseudowords, and pictures of existent and nonexistent objects as either "real" or "made-up." None of the words or pictures corresponded to the same semantic concepts. On each trial, participants were first shown a fixation cross (+) for 500 ms and then were shown a stimulus and were asked to indicate whether it was a real word or object or a made-up object. All stimuli were shown for 500 ms. Then, participants were shown a mask (a block of # symbols) for 250 ms prior to being given 3,000 ms to respond by pressing the Q (real) or P (made-up) keys to indicate their response. Each trial advanced to the next after the participant's response. The 500 ms exposure time was kept constant across all the encoding trials, ensuring that all participants saw all stimuli for the same amount of time. There was a total of 64 encoding trials, evenly divided among real and made-up words and objects (i.e., 16 of each of the four stimulus types). Stimulus presentation order was randomized for each participant. Participants were not warned of a subsequent test phase.

After the encoding phase and an approximately four-minute math distractor task (to minimize the risk of stimulus rehearsal in working memory between the encoding and recognition phases), participants were given a surprise recognition memory test in which they were asked to view more stimuli and decide whether they had just seen them during the encoding phase. The test consisted of 128 items, half of which were presented in the encoding phase (“Old” items) and the other half of which had not been presented in the study (“New” items). “Old” and “New” items consisted of an even distribution of real and made-up words and nonwords. In each recognition phase trial, participants were first shown a fixation for 500 ms and then were shown a stimulus and asked to indicate whether it was “Old” or “New” and press the corresponding keys to enter their response (i.e., Q = old, P = new). Participants were given 5,000 ms to respond on each trial, and each trial advanced to the next after their response. Stimulus presentation order was randomized for each participant.

Other Declarative Memory Tasks

Although recognition memory tasks such as the one just described are canonical measures of declarative memory abilities, we wanted to be confident that the task was ecologically valid. We did so by examining whether it was correlated with commonly used nonverbal (Continuous Visual Memory Test: Trahan & Larrabee, 1988) and verbal declarative memory tasks (MLAT-V: Carroll & Sapon, 1959). The Continuous Visual Memory Test (CVMT: Trahan & Larrabee, 1988) has been employed in several studies assessing individual differences in declarative memory and language (e.g., Hamrick et al., 2019; Morgan-Short et al., 2014; Murphy et al., 2021; Ruiz et al., 2018, 2021; Walker et al., 2020; Zhang et al., 2021). The CVMT examines continuous recognition memory for abstract shapes. In the task, participants were visually presented with a series of complex abstract visual designs (i.e., the visual stimuli do not depict real objects, actions, or concepts) on a computer screen. Seven of the visual designs were repeated throughout the task, with all exposures to these designs after their first presentation constituting the “old” trials, while the other 63 designs were only presented once each and constituted the “new” trials. Each trial presented a visual design for two seconds, followed by a prompt: “New or Old?” The two-second exposure time was kept constant across all the trials, ensuring that all participants saw the designs for the same amount of time. Participants were given one second to respond to the “New or Old?” prompt (i.e., pressing the left-arrow key for “New” and the right arrow for “Old”). The same fixed trial order was given to all participants.

The MLAT-V (Carroll & Sapon, 1959) paired-associate learning task assesses verbal associative learning in declarative memory, and it has been widely used in L2 research (e.g., Buffington & Morgan-Short, 2018; Morgan-Short et al., 2014; Ruiz et al., 2018). Participants are asked to study a series of word pairs with an immediate posttest. Participants first learned twenty-four Kurdish-English word pairs (e.g., hij-draw) for two minutes and were informed that they would be subsequently tested. The test phase consisted of twenty-four multiple-choice items, one test item per trial without a time limit. Each test item comprised a target Kurdish word along with five possible answers in English. Participants were asked to press the corresponding key on the keyboard to indicate their answers (e.g., press the “A” key if they thought the “A” alternative presented was paired with the prompt word in the learning phase). The same fixed trial order was given to all participants.

To that end, all participants completed the three declarative memory tasks in a quiet laboratory using the experiment presentation software PsychoPy (Peirce et al., 2019) in the following order: (1) MLAT-V, (2) our novel declarative memory task, and (3) CVMT.

Data analysis

Data from one subject were removed from all analyses for not paying attention (i.e., playing on their phone) during the encoding phase of our novel declarative memory task (remaining n = 53). Data from 18 participants in the CVMT (remaining n = 35) and data from seven participants in the MLAT-V were removed (remaining n = 46) due to participants not paying attention during the task (i.e., playing on their phones), software problems, or data entry errors. All remaining data were analyzed in R v4.1.3 (R Core Team, 2022). Data and code for all analyses reported here are available via the Open Science Framework (https://osf.io/62exj/?view_only=f17c970770fa4574bc4c4c9dc8f2d78c).

Results

As expected, accuracy in the incidental encoding phase was high (M = 86.77%, SD = 16.71%), with relatively fast reaction times (reaction times (RTs): M = 564 ms, SD = 215 ms). Accuracy was high in the recognition test phase as well, with descriptively similar performance for both verbal (M = 79.48%, SD = 18.81%) and nonverbal (M = 79.92%, SD = 17.47%) items. RTs were slower in the recognition phase than the encoding phase for both verbal (M = 1,113 ms, SD = 237 ms) and nonverbal (M = 1,116 ms, SD = 272 ms) stimuli.

In order to examine the degree to which verbal and nonverbal declarative memory abilities shared the same underlying mechanisms, we sought to estimate the magnitude of the correlations between performance on verbal and nonverbal trials. Since correlations index covariance between variables, they can act as estimates of how much any two variables (in this case, verbal and nonverbal declarative memory abilities) reflect common underlying mechanisms or representations. We computed split-half correlations between participants' responses to verbal and nonverbal stimuli within each phase of the task (encoding and recognition), separately for both stimulus types (real and made-up), for both recognition types (old and new), and for both accuracy and RTs. Split-half correlations take advantage of the utility of resampling data to offset the effect of influential cases in the data (Hamrick, 2019). The split-half correlations and split-half reliabilities were computed over 1,000 resamples of the data for each correlation. The results of the split-half correlation analyses are shown in Table 3. Overall task split-half reliability was high in both the incidental encoding phase (accuracy = .913, RT = .911) and the recognition phase (accuracy = .764, RT = .947).

Although interpretation of correlation coefficients as small, medium, and large is necessarily dependent upon the methodological choices and field of inquiry under consideration, the correlations in performance between verbal and nonverbal stimuli reported in Table 3 are generally considered within the field (Plonsky & Oswald, 2014) to be medium (three coefficients) and large (nine coefficients). These correlations indicate that verbal and nonverbal declarative memory abilities share much underlying variance. To examine this more directly, we conducted regression analyses (separately for accuracy and RT and separately for the encoding and recognition phases) with nonverbal declarative memory abilities as a predictor variable and verbal declarative memory abilities as an outcome variable. Nonverbal declarative memory abilities were a statistically significant predictor of verbal declarative memory abilities in the encoding phase, both for accuracy, $B = 0.77$, $t = 7.73$, $p <. 001$, adjusted $R^2 = .53$, and RT, $B = 0.77$, $t = 8.70$, $p <. 001$, adjusted $R^2 = .59$, as well as in the recognition phase, both for accuracy, $B = 0.65$, $t = 5.62$, $p <. 001$, adjusted $R^2 = .37$, and RT, $B = 0.82$, $t = 13.94$, $p <. 001$, adjusted $R^2 = .79$. This range of R^2 values for

Table 3. Mean split-half correlations and reliabilities between subject-level accuracy and reaction time (RT) performance on verbal and nonverbal stimuli the declarative memory task

	Encoding		Recognition			
			Old		New	
	Real	Made-up	Real	Made-up	Real	Made-up
Mean split-half correlations						
Accuracy	0.803	0.711	0.434	0.601	0.461	0.488
RT	0.712	0.741	0.666	0.753	0.702	0.689
Mean split-half reliabilities						
Accuracy	0.891	0.831	0.605	0.751	0.631	0.656
RT	0.832	0.851	0.799	0.859	0.825	0.816

Table 4. Pearson correlation coefficients (95% confidence intervals in brackets) between accuracy on the verbal-nonverbal declarative memory task and accuracy on either the Continuous Visual Memory Task or the paired-associate learning task

Declarative memory task			
Phase	Verbalness	Continuous Visual Memory Task	Paired-associate learning task
Encoding	Verbal	.531 [.24,. 73], p =. 001	.287 [–.00,. 53], p =. 053
	Nonverbal	.493 [.19,. 71], p =. 002	.313 [.02,. 55], p =. 034
	Overall	.534 [.24,. 74], p <. 001	.316 [.02,. 56], p =. 032
Recognition	Verbal	.402 [.08,. 65], p =. 019	.459 [.19,. 66], p <. 001
	Nonverbal	.413 [.09,. 66], p =. 014	.568 [.33,. 74], p <. 001
	Overall	.451 [.13,. 68], p =. 006	.567 [.33,. 74], p <. 001

these regression models (as well as the R^2 values of all the correlation coefficients in Table 3, which range from 19% to 79% shared variance) can be taken as evidence of partially overlapping to heavily overlapping underlying mechanisms (though we acknowledge that this is qualitative interpretation of the quantitative results that is open for debate). The results also revealed descriptively larger correlations between verbal and nonverbal stimuli for RTs than for accuracy in all but performance at encoding for real stimuli. Correlations were generally smaller for both accuracy and RT in the recognition phase than the incidental encoding phase, with RTs to verbal and nonverbal items still being more strongly correlated than accuracy for verbal and nonverbal items.

To evaluate the degree to which our verbal-nonverbal declarative memory task was correlated with commonly used measures of verbal and nonverbal declarative memory, we examined correlations between accuracy performance on verbal and nonverbal stimuli in the verbal-nonverbal declarative memory task and accuracy on the CVMT and MLAT-V (Table 4). The results indicate a mix of small, medium, and large correlations between our novel declarative memory task and the CVMT and MLAT-V. Correlations between accuracy on either the verbal or nonverbal parts of our declarative memory task were not consistently more correlated with the verbal MLAT-V or the nonverbal CVMT, respectively. Performance in the incidental encoding phase of our declarative memory task was descriptively more correlated with the CVMT than the MLAT-V, while performance in the recognition phase of our declarative memory task was descriptively more correlated with the paired-associate learning task than the CVMT. The correlation between accuracy on the CVMT and the MLAT-V was marginally nonsignificant, $r = .323$, 95% CI [-0.02, 0.60], $p = .067$, but fell within the range reported in previous studies (Buffington et al., 2021; Morgan-Short et al., 2014; Buffington & Morgan-Short, 2018).

Discussion

The aim of this study was to attempt to estimate the size of the relationship between verbal and nonverbal declarative memory abilities while controlling for task demands and stimulus modality. Out of 12 correlations computed, nine coefficients were large, and three were medium by the standards of the field (Plonsky & Oswald, 2014). Inasmuch as correlation coefficients indicate shared variance (e.g., through transformation to R^2 values or as regression model fits), these results indicate that verbal and nonverbal declarative memory abilities may rely on partially-to-heavily overlapping underlying mechanisms. Although correlation values may index overlapping mechanisms or overlapping representations, we argue here that, because the verbal and nonverbal stimuli were not overlapping semantically (e.g., "cone" was presented verbally but not as a nonverbal picture stimulus; a picture of a newspaper was presented as a nonverbal picture stimulus but not as a verbal stimulus), the correlations reported here likely derive not from shared semantic representations of words and objects but rather from similar mechanisms of encoding and retrieval for both verbal and nonverbal stimuli.[2]

Although the present results are broadly consistent with the conclusion that verbal and nonverbal declarative memory abilities in this task tap largely the same underlying mechanisms, the results are more nuanced in the following ways. First, correlations between verbal and nonverbal performance were generally larger for encoding (mean split-half correlation at encoding across both accuracy and RT = .741) than recognition (mean split-half correlation at recognition across both accuracy and RT = .599). What explains this difference? One possibility is that the incidental encoding phase taps encoding and decision-making in a stable, consolidated semantic memory, a subset of the declarative memory system (Ullman, 2016). That is, the encoding phase was, in essence, both an object and lexical decision task, both of which are assumed to tap semantic memory. Although researchers may be tempted to think that these larger correlations within the encoding phase mean that semantic memory tasks might be preferable measures of declarative memory abilities, it is worth noting that such semantic memory tasks typically tap existing knowledge (e.g., Murphy et al., 2021), which may not necessarily correlate with recently acquired L2 knowledge if that knowledge is represented in episodic memory (e.g., Witzel & Forster, 2012; Hamrick et al., 2019), which relies on a partially distinct neural substrate of declarative memory (Ullman, 2016; Davis & Gaskell, 2009).

In comparison, the smaller correlations between verbal and nonverbal performance at recognition are consistent with the idea that recognition more strongly taps episodic memory. Episodic memory representations may widely vary in their representational strength, which could add noise to the data, lowering the size of the obtained correlation coefficients. Moreover, episodic memory performance can be supported by distinct declarative mechanisms of recollection or familiarity (Yonelinas, 2002). Subject- or item-level variability in reliance on these different declarative memory mechanisms

---

[2] However, this does not mean that words and objects do not share some level of representation. Indeed, a large body of evidence has converged on the view that word and object semantic representations are either overlapping or identical (e.g., Evans, 2016).

may also have added noise to the data, lowering the obtained correlations. If true, this account of our findings would indicate that declarative memory tasks that rely on episodic memory performance to operationalize declarative memory abilities may be tapping distinct, albeit related, declarative mechanisms, adding to noise in the data, and consequently leading researchers to underestimate the actual size of the true relationship between declarative memory abilities and any recently learned language information, assuming the latter would be episodically represented as well, which recent research has shown to be a very likely possibility (e.g., Hamrick et al., 2019; Murphy et al., 2021; Witzel & Forster, 2012; Zhang et al., 2021). However, more research is needed to address this possibility.

Second, performance on our declarative memory task was significantly correlated (range of *r*s = .313–.568) with performance on both the CVMT and the MLAT-V in all but one comparison (*r* = .287), indicating some evidence of convergent validity with commonly used declarative memory tasks. Our novel verbal-nonverbal declarative memory task was more correlated with performance in the CVMT (*r* = .471) than the MLAT-V (*r* = .418) on average, as would be expected given that our task and the CVMT are both recognition memory tasks. Although some correlations were numerically larger between our task and the MLAT-V, none of these differences in correlation size were statistically significant, *p*s >. 36. However, more research should be conducted to determine the correlations between verbal and nonverbal implementations of recognition memory and associative learning tasks, since the pattern of results with our task, the CVMT, and the MLAT-V did not neatly align with what might be expected in accordance with their task demands or verbalness.

Implications for L2 research

As individual differences research on memory and L2 learning rapidly expands, it is critical to bear in mind that many of the cognitive constructs we are interested in (e.g., declarative memory) are complex phenomena that are not monolithic or homogenous and, therefore, cannot be exhaustively measured in a single task. Since any single task can only tell us about a part of a construct, it is important to be cautious not to overstate or overgeneralize the association between that construct and L2 learning. This is clearly true in the study of individual differences in declarative memory abilities. There are multiple subsystems of declarative memory, and there are multiple mechanisms implicated within and across those subsystems. Consequently, different tasks may tap different mechanisms and aspects of the larger declarative system. Bearing in mind that recognition memory tasks (like the one reported here) are a popular way to assess declarative abilities, it is worth considering that recognition accuracy results showed the smallest correlations between performance verbal and nonverbal stimuli of all the analyses we conducted. Thus, we encourage future L2 researchers who seek to use recognition accuracy as a measure of individual differences in declarative memory abilities to do so with a nonverbal measure. If verbal and nonverbal recognition accuracy abilities were strongly correlated enough to conclude that they were more or less isomorphic, then the choice of verbal versus nonverbal stimuli might not matter. However, given their relatively smaller correlations (from the array of correlation magnitudes that we found), it is worth exercising caution and not concluding such an isomorphic relationship. This underscores the importance of using nonverbal memory measures when conducting individual differences research in L2 learning in order to avoid the problem of shared verbalness (Morgan-Short et al., 2022).

For the time being, we recommend that researchers use multiple measures of declarative memory abilities when possible, and if it is not possible, they should (a) carefully consider which type of declarative memory task makes the most sense for addressing their research question and (b) interpret the findings in terms of the specific declarative abilities tapped by the task rather than drawing conclusions about the declarative memory system broadly. The small- to medium-sized correlations between our verbal-nonverbal declarative memory task, the MLAT-V, and the CVMT underscore this point. Researchers should also consider which declarative mechanisms are relevant for their own research questions. If associative learning is key, then a nonverbal paired-associates analog to the MLAT-V might prove ideal. If recognition is key, then tasks like the CVMT or other recognition memory tasks (e.g., Lukácz et al., 2017) may prove more useful.

Limitations and future directions

Caution is warranted in interpreting the current findings for several reasons. First, having unequal sample sizes for different tasks may have contributed to reduced power to detect a significant correlation for the one nonsignificant correlation in our data (between performance verbal encoding in our declarative memory task and performance on the MLAT-V). Analyses were also limited by the fact that some of the accuracy data from our task were skewed, which could have reduced the overall effect size estimate. Future research with a more challenging version of this task (i.e., to induce individual differences and avoid ceiling effects) should be carried out to replicate and extend this work. Additionally, our study may also have been affected by the use of the same response keys (Q and P on a computer keyboard for both the encoding and the recognition phases). Perhaps most obviously, our declarative memory task is limited in that it operationalized declarative memory abilities via only lexical/object decision and recognition memory. As noted above, paired-associate learning tasks, such as the MLAT-V and other associative learning tasks, may recruit different declarative memory substrates. As a consequence, it is unknown whether the correlations observed between verbal and nonverbal abilities in our declarative memory task would be similar to correlations that would be observed in tasks that encourage the use of associative learning mechanisms in declarative memory. Our study was also limited in that it was not preregistered.

Another limitation of our study is that the correlations between performance on verbal and nonverbal stimuli in our declarative memory task may be due to shared factors that are *not* underpinned by declarative memory abilities, per se. For example, one could argue that attentional resources and decision-making strategies might have been similar between verbal and nonverbal stimuli, and, hence, the resulting correlations might have been due, in part or whole, to these other cognitive abilities. However, we find such a strong objection to be premature for two reasons. First, the recognition memory tasks are well-understood measures of declarative memory, both at the psychological and neurological level, which should bolster our confidence that the current recognition memory paradigm was, in fact, a valid measure of declarative memory. Second, the possibility that shared task demands (e.g., attention, decision-making, as mentioned above) could confound the results of any correlational individual differences study. In fact, it is one of the chief limitations of such research, hence the need to look for converging evidence across a range of methodologies, as noted in other recent papers (e.g., Hamrick et al., 2019).

Finally, it is notable that the correlations between verbal and nonverbal RT performance were larger than those between verbal and nonverbal accuracy performance. One simple explanation for this finding is that although RT distributions deviated only mildly from a Gaussian distribution, the distribution of accuracy scores was high and negatively skewed, producing smaller correlation coefficients. However, there are other plausible possibilities. For example, our findings could be interpreted in terms of the diffusion model of two-choice decision tasks (Ratcliff, 1978). Applied to our data, the diffusion model would propose that RTs were a byproduct of a self-terminating decision process based on noisy evidence accumulation. Higher correlations between RTs for verbal and nonverbal stimuli may simply indicate that the mechanisms of encoding, retrieval, and decision-making are the same for verbal and nonverbal materials, whereas the lower correlations for accuracy may indicate item-level differences in the representation of verbal and nonverbal stimuli. However, this explanation is speculative, and more research is needed to adjudicate between it and the simpler explanation based on skewness.

Conclusion

Declarative memory is theorized to play a critical role in L2 (Ullman, 2016; Witzel & Forster, 2012), and individual differences research has shown that declarative memory abilities contribute to a wide array of L2 phenomena. However, studies reported a wide range of different effect sizes and, in some cases, null results (Hamrick et al., 2018; Morgan-Short et al., 2022), and the use of different declarative memory tasks with different task demands (recognition vs. associative learning) and stimulus properties (verbal vs. nonverbal stimuli) may be partially responsible for such mixed results. Similarly to what Buffington et al. (2021) found for procedural memory research in L2 learning, it is not clear that tasks purported to tap declarative memory all do so in the same way. There are multiple subsystems and mechanisms implicated in declarative memory, and L2 research examining individual differences in declarative memory must take this complexity seriously. Although declarative memory is a useful construct for psychologists, neuroscientists, and linguists, it is not a monolithic system or ability. Until we systematically determine the links between task demands (e.g., recognition vs. encoding vs. associative learning) and stimulus domain (e.g., verbal vs. nonverbal) and how they influence the recruitment of different declarative abilities, caution will need to be taken in how we describe it, theorize about it, and measure it.

The present study sought to examine the relationship between verbal and nonverbal declarative memory abilities within the same modality and within the same task instructions in order to control for stimulus and task demands. In principle, this design should have maximized the correlations between verbal and nonverbal declarative memory abilities since the only difference between the two conditions was the verbal versus nonverbal nature of the stimuli. However, the shared variance between verbal and nonverbal declarative memory abilities ranged from 19% to 79%, indicating that there were both substantial differences and substantial overlap between verbal and nonverbal declarative memory abilities, even after controlling for stimulus and task demands. Until more research is conducted to elucidate the true nature of the relationships between verbal and nonverbal declarative memory abilities, L2 researchers should exercise caution in interpreting their results. If they find, for example, that individual differences in declarative memory abilities do or do not correlate with a given L2 measure, they should consider the effects of the task demands (e.g., associative learning vs. recognition) and stimulus modality (e.g., visual vs. auditory) as sources of noise in the data. They should also consider that even when these variables are controlled for, associations between verbal and nonverbal declarative abilities within the exact same task, although often sizeable, are themselves quite variable. Any successful account of individual differences in declarative memory and L2 learning will critically depend upon addressing both the true size of these associations as well as the sources of variance that moderate them.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S0272263124000093.

Data availability statement. The experiment in this article earned an Open Data badge for transparent practices. The materials are available at https://osf.io/62exj/?view_only=f17c970770fa4574bc4c4c9dc8f2d78c

Competing interest. The author declares no competing interests.

References

Arndt, J. (2012). Paired-associate learning. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 2551–2552). Springer. https://doi.org/10.1007/978-1-4419-1428-6_1038

Artieda, G., & Muñoz, C. (2016). The LLAMA tests and the underlying structure of language aptitude at two levels of foreign language proficiency. *Learning and Individual Differences*, *50*, 42–48. https://doi.org/10.1016/j.lindif.2016.06.023

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*, 445–459. https://doi.org/10.3758/bf03193014

Bowles, A. R., Chang, C. B., & Karuzis, V. P. (2016). Pitch ability as an aptitude for tone learning. *Language Learning*, *66*(4), 774–808. https://doi.org/10.1111/lang.12159

Buffington, J., Demos, A. P., & Morgan-Short, K. (2021). The reliability and validity of procedural memory assessments used in second language acquisition research. *Studies in Second Language Acquisition*, *43*(3), 635–662. https://doi.org/10.1017/s0272263121000127

Buffington, J., & Morgan-Short, K. (2018). Construct validity of procedural memory tasks used in adult-learned language. *Proceedings of the 40th Annual Conference of the Cognitive Science Society*, *1*, 1420–1425. https://www.academia.edu/37235721/Construct_Validity_of_Procedural_Memory_Tasks_Used_in_Adult_Learned_Language

Carroll, J. B., & Sapon, S. M. (1959). *Modern Language Aptitude Test*. The Psychological Corporation/Harcourt Brace Jovanovich. https://psycnet.apa.org/record/1960-03629-000

Davachi, L. (2006). Item, context, and relational episodic encoding in humans. *Current Opinion in Neurobiology*, *16*(6), 693–700. https://doi.org/10.1016/j.conb.2006.10.012

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1536), 3773–3800. https://doi.org/10.1098/rstb.2009.0111

DeKeyser, R. (2020). Input is not a panacea. *International Journal of Bilingualism*, *24*(1), 79–81. https://doi.org/10.1177/1367006918768371

Eichenbaum, H. (2012). *The cognitive neuroscience of memory: An introduction* (2nd ed.). Oxford University Press. https://academic.oup.com/book/12065

Ettlinger, M., Bradlow, A. R., & Wong, P. C. M. (2014). Variability in the learning of complex morphophonology. *Applied Psycholinguistics*, *35*(4), 807–831. https://doi.org/10.1017/S0142716412000586

Evans, V. (2016). Design features for linguistically-mediated meaning construction: The relative roles of the linguistic and conceptual systems in subserving the ideational function of language. *Frontiers in Psychology*, *7*, 156. https://doi.org/10.3389/fpsyg.2016.00156

Faretta-Stutenberg, M., & Morgan-Short, K. (2018). The interplay of individual differences and context of learning in behavioral and neurocognitive second language development. *Second Language Research*, *34*(1), 67–101. https://doi.org/10.1177/0267658316684903

Granena, G. (2019). Cognitive aptitudes and L2 speaking proficiency: Links between LLAMA and HI-LAB. *Studies in Second Language Acquisition*, *41*(2), 313–336. https://doi.org/10.1017/S0272263118000256

Hamrick, P. (2015). Declarative and procedural memory abilities as individual differences in incidental language learning. *Learning and Individual Differences*, *44*, 9–15. https://doi.org/10.1016/j.lindif.2015.10.003

Hamrick, P. (2019). Adjusting regression models for overfitting in second language research. *Journal of Research Design and Statistics in Linguistics and Communication Science*, *5*(1–2), 107–122. https://doi.org/10.1558/jrds.38374

Hamrick, P., Graff, C., & Finch, B. (2019). Contributions of episodic memory to novel word learning. *The Mental Lexicon*, *14*(3), 381–398. https://doi.org/10.1075/ml.19019.ham

Hamrick, P., Lum, J. A. G., & Ullman, M. T. (2018). Child first language and adult second language are both tied to general-purpose learning systems. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 1487–1492. https://doi.org/10.1073/pnas.1713975115

Kroll, J. F., & Potter, M. C. (1984). Recognizing words, pictures, and concepts: A comparison of lexical, object, and reality decisions. *Journal of Verbal Learning & Verbal Behavior*, *23*(1), 39–66. https://doi.org/10.1016/S0022-5371(84)90499-7

Li, S. (2017). An exploratory study on the role of foreign language aptitudes in instructed pragmatics learning in L2 Chinese. *Chinese as a Second Language Research*, *6*(1), 103–128. https://doi.org/10.1515/caslar-2017-0005

Linck, J. A., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., Smith, B. K., Bunting, M. F., & Doughty, C. J. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning*, *63*(3), 530–566. https://doi.org/10.1111/lang.12011

Lukács, A., Kemény, F., Lum, J. A. G., & Ullman, M. T. (2017). Learning and overnight retention in declarative memory in specific language impairment. *PLoS One*, *12*(1), e0169474. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5207735/

Meara, P. (2005). *LLAMA language aptitude tests: The manual.* Lognostics. https://www.llamatests.org/docs/llama_manual.pdf

Morgan-Short, K., Faretta-Stutenberg, M., Brill-Schuetz, K., Carpenter, H., & Wong, P. C. M. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition*, *17*(1), 56–72. https://doi.org/10.1017/S1366728912000715

Morgan-Short, K., Hamrick, P., & Ullman, M. T. (2022). Declarative and procedural memory as predictors of second language development. In S. Li, P. Hiver, & M. Papi (Eds.), *The Routledge handbook of second language acquisition and individual differences* (pp. 67–81). https://www.taylorfrancis.com/chapters/edit/10.4324/9781003270546-6/declarative-procedural-memory-predictors-second-language-development-kara-morgan-short-phillip-hamrick-michael-ullman

Murphy, J., Miller, R., & Hamrick, P. (2021). Contributions of declarative memory and prior knowledge to incidental L2 vocabulary learning. *The Mental Lexicon*, *16*, 49–68. https://doi.org/10.1075/ml.20012.mur

Paradis, M. (2009). *Declarative and procedural determinants of second languages* (vol. *40*). John Benjamins. https://doi.org/10.1075/sibil.40

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*. https://doi.org/10.3758/s13428-018-01193-y

Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, *64*(4), 878–912. https://doi.org/10.1111/lang.12079

Pili-Moss, D., Brill-Schuetz, K. A., Faretta-Stutenberg, M., & Morgan-Short, K. (2020). Contributions of declarative and procedural memory to accuracy and automatization during second language practice. *Bilingualism: Language and Cognition*, *23*(3), 639–651. https://doi.org/10.1017/S1366728919000543

R Core Team (2022). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Rogers, V., Meara, P., & Rogers, B. (2023). Testing language aptitude: LLAMA evolution and refinement. In Z. E. Wen, P. Skehan, & R. Sparks (Eds.), *Language aptitude: Theory and practice* (pp. 47–72). Cambridge University Press. https://doi.org/10.1017/9781009076463.005

Ruiz, S., Tagarelli, K. M., & Rebuschat, P. (2018). Simultaneous acquisition of words and syntax: Effects of exposure condition and declarative memory. *Frontiers in Psychology*, *9*, 1168. https://doi.org/10.3389/fpsyg.2018.01168

Ruiz, S., Rebuschat, P., & Meurers, D. (2021). The effects of working memory and declarative memory on instructed second language vocabulary learning: Insights from intelligent CALL. *Language Teaching Research*, *25*(4), 510–539. https://doi.org/10.1177/1362168819872859

Saito, K. (2017). Effects of sound, vocabulary, and grammar learning aptitude on adult second language speech attainment in foreign language classrooms. *Language Learning*, *67*(3), 665–693. https://doi.org/10.1111/lang.12244

Saito, K. (2019). The role of aptitude in second language segmental learning: The case of Japanese learners' English /ɹ/ pronunciation attainment in classroom settings. *Applied Psycholinguistics*, *40*(1), 183–204. https://doi.org/10.1017/S0142716418000528

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174–215. https://doi.org/10.1037/0278-7393.6.2.174

Trahan, D. E., & Larrabee, G. J. (1988). *Continuous visual memory test: Professional manual.* Psychological Assessment Resources. https://www.parinc.com/products/pkey/78

Ullman, M. T. (2004). Contribution of memory circuits to language: The declarative/procedural model. *Cognition*, *92*, 231–270. https://doi.org/10.1016/j.cognition.2003.10.008

Ullman, M. T. (2016). The declarative/procedural model: A neurobiological model of language learning, knowledge, and use. In G. Hickok & S. A. Small (Eds.), *The neurobiology of language* (pp. 953–968). Elsevier. https://doi.org/10.1016/B978-0-12-407794-2.00076-6

Ullman, M. T. (2020). The declarative/procedural model: A neurobiologically-motivated theory of first and second language. In B. VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second language acquisition* (3rd ed., pp. 128–161). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9780203628942-13/declarative-procedural-model-neurobiologically-motivated-theory-first-second-language-michael-ullman

Walker, N., Monaghan, P., Schoetensack, C., & Rebuschat, P. (2020). Distinctions in the acquisition of vocabulary and grammar: An individual differences approach. *Language Learning*, *70*(S2), 221–254. https://doi.org/10.1111/lang.12395

Witzel, N. O., & Forster, K. I. (2012). How L2 words are stored: The episodic L2 hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(6), 1608. https://doi.org/10.1037/a0028072

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Yonelinas, A. P., Widaman, K., Mungas, D., Reed, B., Weiner, M. W., & Chui, H. C. (2007). Memory in the aging brain: Doubly dissociating the contribution of the hippocampus and entorhinal cortex. *Hippocampus*, *17*(11), 1134–1140. https://doi.org/10.1002/hipo.20341

Zhang, Y., Ridchenko, M., Hayashi, A., & Hamrick, P. (2021). Episodic memory contributions to second language lexical development persist at higher proficiencies. *Applied Cognitive Psychology,* 1–6. https://doi.org/10.1002/acp.3865

Cite this article: Hamrick, P., Was, C. A., & Zhang, Y. (2024). Do verbal and nonverbal declarative memory tasks in second language research measure the same abilities? *Studies in Second Language Acquisition,* 46: 581–596. https://doi.org/10.1017/S0272263124000093