Word Meaning Contributes to Free Recall Performance in Supraspan Verbal List-Learning Tests

Sandrine Cremona*, Gaël Jobard, Laure Zago and Emmanuel Mellet

Groupe d'Imagerie Neurofonctionnelle, IMN, UMR 5293, CNRS, Univ. Bordeaux, CEA, Bordeaux, France

Supraspan verbal list-learning tests, such as the Rey Auditory Verbal Learning Test (RAVLT), are classic neuropsychological tests for assessing verbal memory. In this study, we investigated the impact of the meaning of the words to be learned on three memory stages [short-term recall (STR), learning, and delayed recall (DR)] in a cohort of 447 healthy adults. First, we compared scores obtained from the RAVLT (word condition) to those of an alternative version of this test using phonologically similar but meaningless items (pseudoword condition) and observed how each score varied as a function of age and sex. Then, we collected the participants' self-reported strategies to retain the word and pseudoword lists and examined if these strategies mediated the age and sex effects on memory scores. The word condition resulted in higher memory scores than pseudoword condition at each memory stage and even canceled out, for the learning stage, the detrimental effect of age that was observed for the short-term and DR. When taking sex into account, the word advantage was observed only in women for STR. The self-reported strategies, which were similar for words and pseudowords, were based on the position of the item on the list (word: 53%, pseudoword: 37%) or the meaning of the item (word: 64%, pseudoword: 58%) and were used alone or in combination. The best memory performance was associated with the meaning strategy in the word condition and with the combination of the meaning and position strategies in the pseudoword condition. Finally, we found that the word advantage observed in women for STR was mediated by the use of the meaning strategy. The RAVLT scores were thus highly dependent on word meaning, notably because it allowed efficient semantic knowledge-based strategies. Within the framework of Tulving's declarative memory model, these results are at odds with the depiction of the RAVLT as a verbal *episodic memory* test as it is increasingly referred to in the literature.

Keywords: Rey Auditory Verbal Learning Test (RAVLT), semantic memory, self-reported strategies, adult healthy human volunteers, BIL&GIN

INTRODUCTION

Memory is among the cognitive functions that change the most with aging. With the current aging of the population, the study of memory during normal and pathological aging has become a major focus in the neuroscience field (Park and Festini, 2017). Neuropsychological assessment of learning and memory in aging populations frequently relies on the use of learning tasks based on supraspan word-lists (i.e., list lengths that exceed memory span), such as those from the Consortium to Establish a Registry for Alzheimer's Disease (CERAD; Morris et al., 1989), the California Verbal Learning Test (CVLT; Delis et al., 1987) or the Rey Auditory Verbal Learning Test (RAVLT; Rey, 1970). These tests include measures of short- and long-term recall and recognition, vulnerability to proactive and retroactive interference and learning ability. Although the CVLT yields more information, particularly regarding the encoding strategies used by the participants (Stricker et al., 2002), the RAVLT has long offered more adequate normative data and alternative test forms (Spreen et al., 1998). As reviewed by Saury and Emanuelson (2017), the extended version of the RAVLT can distinguish among the following four domains of learning and memory: (a) short-term recall (STR), (b) learning, (c) interference, and (d) retrieval [immediate recall after interference, delayed recall (DR), and delayed recognition]. Thus, the RAVLT provides numerous memory scores and derivative indices (Spreen et al., 1998). At the neuropsychological level, RAVLT performance (STR, learning, and DR) can discriminate among normal aging, mild cognitive impairment (MCI) and Alzheimer's disease (AD; Marra et al., 2000; Estévez-González et al., 2003; Goryawala et al., 2015; Bauer et al., 2018) and is a good independent predictor of dementia (Eckerström et al., 2013; Li et al., 2017). Moreover, RAVLT DR performance can discriminate AD from other dementias, such as the behavioral variant of frontotemporal dementia (Ricci et al., 2012) or dementia with Lewy bodies (Bussè et al., 2017). Therefore, the clinical use of the RAVLT to detect memory impairment is supported by the French social security system (ALQP006 CCAM). Taken together, these elements can explain why the RAVLT is currently highly popular as shown by its inclusion in cognitive batteries used for aging studies [e.g., the Alzheimer's Disease Neuroimaging Initiative (ADNI, Park et al., 2012) and the Canadian Longitudinal Study on Aging (CLSA, Tuokko et al., 2017)].

The present study aimed to further characterize the role of verbal material, i.e., its meaning, in memory assessment using supraspan verbal list-learning tests such as the RAVLT. To achieve our goal, we adopted a two-pronged approach that involved a healthy adult population of 447 participants under the age of 60 to avoid the pronounced effects of aging on memory.

First, to study the role of the lexico-semantic component of the verbal material, we compared memory scores obtained from the RAVLT to those of an alternative version of this test using phonologically similar but meaningless items, i.e., pseudowords. To the best of our knowledge, this issue has not been explored in this manner since certain preliminary works were carried out by Ebbinghaus on himself and by Tulving on a group of six participants (Tulving, 1985). Pseudowords are orthographically legal and pronounceable letter strings without meaning that should mainly activate orthographic and phonological sources of information but only partially (or not at all) lexical information (Mazoyer et al., 1993; Ziegler et al., 1997). Therefore, we were able to evaluate the lexico-semantic benefit by comparing pseudoword from word memory performance. As a first assumption, we expected that word meaning would enable better word than pseudoword memory performance. Moreover, demographic factors have been shown to have a major impact on RAVLT scores in the following manner: performance declines during healthy aging (McMinn et al., 1988; Geffen et al., 1990; Selnes et al., 1991; Mortensen and Gade, 1993; Van Der Elst et al., 2005; Malloy-Diniz and Parreira, 2007; Messinis et al., 2007; Teruya et al., 2009), and women perform better than men (Bleecker et al., 1988; Geffen et al., 1990; Mortensen and Gade, 1993; Aartsen et al., 2004; Maitland et al., 2004; Malloy-Diniz and Parreira, 2007; Messinis et al., 2007). We therefore examined the variability of the RAVLT and the pseudoword memory scores as a function of age and sex.

Second, we collected the participants' self-reports of the strategies they used during the task to assess whether the self-reported strategies could lead to a better understanding of the explicit mechanisms by which the learning and memory occur. The use of encoding strategies has been linked to greater recall in word-list learning tests (Unsworth et al., 2019). Compared to the CVLT, in which the words are experimentally grouped into four categories to encourage the use of a semantic clustering strategy (Sundararaman et al., 2013), the RAVLT words do not have a clear semantic relationship. Some authors have argued that serial order might become a preferred strategy in such cases (Vakil and Blachstein, 1994; Meijs et al., 2013). Here, we investigated the nature of the participants' declarative memory processes without any preconception by collecting their self-reported strategies. Therefore, our second hypothesis was that RAVLT performance could be influenced by the self-reported strategies that the participants used. We also investigated whether these strategies could mediate the well-known effect of age and sex on RAVLT performance.

MATERIALS AND METHODS

Participants

The present study included 447 healthy volunteers recruited during the same period whose native language tongue was French and who had no past history of neuropsychiatric disorders. The data were derived from the BIL&GIN database (Mazoyer et al., 2016), and we excluded four participants with unavailable data and two participants with suspected dyslexia. The experimental protocol was approved by the local ethics committee (Comité de Protection des Personnes Nord-Ouest). All participants provided written informed consent and received compensation for their participation. The study sample comprised 228 women and 219 men with a mean age of 26.6 years (range from 18 to 58 years) and was balanced for handedness. The mean level of education was 15.3 years, which corresponds to 3 years of university education.

Cognitive Tests

Data were extracted from the standardized battery of 10 cognitive tests in the BIL&GIN, a database acquired by the team and previously described in detail (Mazoyer et al., 2016). Briefly, the participants completed 10 cognitive tests distributed over two sessions separated by an approximately 3-h MRI acquisition period. We used six verbal tests that were ordered within each cognitive session as follows:

– Session 1: the RAVLT, vocabulary scope, a rhyme judgment task and the listening span test.
– Session 2: the pseudoword memory test and the reading span test.

Assessment of the Memory of Words (RAVLT) and Pseudowords

Since the RAVLT has severe ceiling effects in healthy young adults (i.e., those 18–39 years of age, Uttl, 2005), we attempted to limit the ceiling effect by increasing the number of words to be learned. In total, 18 words were used (livre, fleur, train, tapis, prairie, harpe, sel, jardin, doigt, tambour, pomme, cheminée, rivière, bouton, clé, chien, verre, and hochet, Rey, 1970; Lezak, 1983) instead of the 15 words used in the common version. The verbal memory test consisted of listening to a list of concrete "unrelated" words (one word per second read aloud by the experimenter in the same order in each of the five consecutive trials). Immediately after each trial, the participants were instructed to freely recall as many items as possible. After a 20-min delay (during which a nonverbal task was performed), the participants were asked to recall this list again. A similar procedure was applied using 15 pseudowords created with the WordGen software (Duyck et al., 2004) and matched to the list of words by the number of letters, phonemes, syllables and bigram frequencies (guice, anire, ficot, meple, flaxion, jaron, asue, ecrot, diare, doussant, boidir, sato, fince, veigne, and gouage).

As the measures were collected in a healthy adult population, we expected a relatively high level of DR (Van Der Elst et al., 2005). Additionally, since recognition requires fewer processing resources than recall (Craik and McDowd, 1987), we did not include a recognition task and focused instead on three scores (Ivnik et al., 1990) corresponding to the following three stages of memory:

1. STR: recall score in trial 1;
2. Learning over trials (LOT): the sum of the items recalled over trials 2 & 3 minus 2 times the STR (Ivnik et al., 1990; Teruya et al., 2009); and
3. DR: the long-term percent retention was calculated as the delayed trial six score divided by the maximum score achieved during one of the first five learning trials × 100 (Moradi et al., 2017).

The detailed method for determining the most suitable scores for evaluating each memory stage is described in the **Supplementary Material**. To avoid any ceiling effect, Uttl advised that the mean of each score be distanced from the maximum score by more than one SD (Uttl, 2005). Thus, we checked that all mean memory scores met this criterion.

Assessment of Confounding Factors

Vocabulary scope

Greater vocabulary and verbal IQ have been associated with better RAVLT performance (Bolla-Wilson and Bleecker, 1986; Mortensen and Gade, 1993). A synonym-finding test served to estimate the extent of vocabulary (Binois and Pichot, 1956). Across 44 trials (max score), the participants had to determine which of the six written words was synonymous with a target word presented at the top of the screen.

Rhyme judgment task (rhyming)

Phonological abilities could impact pseudoword processing. To estimate phonological abilities, a rhyme judgment task (adapted from Shaywitz et al., 1995) was completed using 80 pairs of pseudowords. The pairs of pseudowords were presented and remained on the screen until the participant indicated by pressing a key whether or not they rhymed or for a maximum of 4 s. The pseudowords were composed of 1–3 syllables (3–9 letters); 30 pairs rhymed, while 50 did not, leading to a maximum possible score of 80. To prevent the use of strategies based simply on visually matching between the ends of the pseudowords, 45 items were constructed so that the visual information conflicted with the expected responses (i.e., by using pseudowords with similar spellings but different sounds or pseudowords with different spellings but same sounds).

Working memory capacity

Working memory capacity (WMC) has been shown to mediate the rate of age-related decline in verbal memory (Constantinidou et al., 2014). Here, WMC was evaluated with two complex working memory span tasks: the French adaptation of the reading span test and its auditory counterpart, the listening span test (Daneman and Carpenter, 1980; Desmette et al., 1995). During the reading span test, the participant had to read aloud blocks of sentences presented on a computer screen. The number of sentences per block started at 2 and increased by 1 sentence every 3 blocks until 6 sentences were presented. At the end of each block, the participant had to remember the last word of each sentence while avoiding starting with the last sentence. The same pattern was used for the listening span test but with 2 exceptions, as follows: each sentence was read by the examiner, and the participants had to determine whether the sentence was in the present tense. According to the Daneman and Carpenter method, truncated spans were scored for each test by starting with the highest level (2–6) at which the participant recalled the majority of the blocks (2 out of 3) and adding half a point for recalling 1 out of 3 blocks at the subsequent level. The truncated spans for reading and listening were averaged to obtain the WMC score.

Vocabulary scope, rhyming task and reading span test were conducted with E-Prime (Version 2, Pittsburgh, PA: Psychology Software Tools.).

Collection and Categorization of Self-Reported Strategies

Self-reported strategies were collected after the word and pseudoword DR tasks by a psychologist trained in interviewing. The self-reported strategies corresponded to all three memory

stages (STR, LOT, and DR). The participants were asked to explain how they retained the lists, and their answers were recorded as field notes. An open question was used to avoid influencing the nature of the answers. At this stage, data related to words from three participants were missing, as were data related to pseudowords from five participants.

Based on the psychologist's field notes, two experimenters compiled an exhaustive list of the strategies used for the words and pseudowords and standardized the wording of the strategies on the basis of consensus. This standardization was facilitated by the fact that the psychologist's field notes were already partially standardized. A set of categories that covered all strategies and allowed word and pseudoword comparisons was obtained according to the following rules (**Table 1**):

1. Since an auditory representation of the items was directly induced by the auditory presentation of the lists, Listening was the default strategy.
2. If a strategy involved rethinking the order of the list, it was considered a Position strategy.
3. If a strategy implied knowledge about the meaning of the items (words) or a phoneme-to-semantic system conversion (pseudowords), it was considered a Meaning strategy.
4. If a strategy referred to the participant's time-space, it was considered an Autonoetic strategy.

Finally, the participants were categorized into four types of strategy use as follows:

1. Those who only used Listening strategies (words $n = 67$; pseudowords $n = 110$);
2. Those who only used Position strategies (words $n = 93$; pseudowords $n = 80$);
3. Those who only used Meaning strategies (words $n = 142$; pseudowords $n = 168$);
4. Those who used both Position and Meaning strategies, which was labelled Dual (words $n = 142$; pseudowords $n = 84$).

The Listening strategy that was directly induced by the auditory presentation of the items was regarded as a *minimal* strategy compared to the other 3, which were considered more *advanced*. Moreover, since autonoetic strategies were used by only 3% of the participants and were never used alone, except by one participant, the Autonoetic category was not separately addressed as a strategy type in further analysis.

Data Analysis

The statistical analyses were conducted in R Version 3.5.2.[1] The figures were created with the following packages: *ggplot2* (Wickham, 2016) and *interactions* (Long, 2019). The alpha level of 0.05 or the 95% confidence interval (CI) was used to determine the significance of the hypothesis tests.

---

[1]http://www.R-project.org

TABLE 1 | Frequency of use and categorization criteria for the set of standardized self-reported word and pseudoword strategies.

Words	N	Involves an auditory representation	Involves rethinking the order of the list	Implies knowledge about the meaning of the items	Refers to the participants' time-space	Categorization of the words and pseudowords self-reported strategies	Refers to the participants' time-space	Implies a phoneme-to-semantic system conversion	Involves rethinking the order of the list	Involves an auditory representation	N	Pseudowords
No/few strategy	10	X				Listening				X	10	No/few strategy
(only) Listening	79	X								X	70	(only) Listening
Rehearsal	31	X								X	26	Rehearsal
Sound/sonorities	31	X								X	56	Sound/sonorities
Rhymes	9	X								X	2	Rhymes
Ordered rehearsal	5	X	X			Position			X	X	6	Ordered rehearsal
Bundled	15	X	X						X	X	9	Bundled
Ordered, by rote	146	X	X						X	X	107	Ordered, by rote
First/beginning	103	X	X						X	X	50	First/beginning
Last/end	75	X	X						X	X	43	Last/end
Taxonomic categorization	87	X		X		Meaning		X		X	251	Association with real/known words
Thematic relation	130			X								
Image/visualization	67			X								
Story	9			X								
Recollection	11	X			X	Autonoetic	X			X	1	Recollection
Test room item	3	X			X		X			X	0	Test room item

Group Comparisons of the Word and Pseudoword Memory Scores

Each memory score (STR, LOT, and DR) was averaged across participants, and the mean scores for the words and pseudowords were compared with a paired sample t-test Bonferroni corrected for the multiplicity of tests ($p < 0.016$). Each individual memory score was further standardized, thereby allowing a direct comparison in the subsequent analyses.

Modeling Memory Performance by Age and Sex

To estimate the effects of age and sex on each memory stage of words and pseudowords, we implemented a linear mixed model (*lmerTest::lmer*, Kuznetsova et al., 2017). A 4-way interaction among age * sex * type of item (2 levels: words and pseudowords) * memory stage (3 levels: STR, LOT, and DR) was introduced in the model as a fixed effect. We further included the following variables as confounders: vocabulary scope, rhyming, WMC, level of education (Selnes et al., 1991; Van Der Elst et al., 2005; Malloy-Diniz and Parreira, 2007; Messinis et al., 2007; Teruya et al., 2009; Magalhães and Hamdan, 2010), and handedness (Mellet et al., 2013). The random effects of type of item and memory stage were fitted on the intercept at the participant level. A stepwise backward strategy for model selection based on the Akaike information criterion (AIC) was applied to the linear mixed model described above (*lmerTest::step*, Kuznetsova et al., 2017) to determine the best-fitted model. The ANOVA components were calculated based on the final model with the Kenward-Roger approximation (Halekoh and Højsgaard, 2014) to correct for the underestimation of variance due to sampling fluctuations. The residuals were visually inspected to assess normality and homoscedasticity. The proportion of variance explained by the fixed effects and their interactions was estimated with the marginal R^2, as described in Nakagawa et al. (2017). All pairwise *post hoc* comparisons were corrected for multiple comparisons using an FDR-controlling procedure (*emmeans*, Lenth, 2020).

Analyses of Self-Reported Strategies

The distributions of the self-reported strategies, which were separately reported for the pseudowords and words, were compared with the test of marginal homogeneity, an extension of McNemar's test for dependent samples for multilevel variables. For each strategy, pairwise *post hoc* comparisons between the word and pseudoword proportions were performed with McNemar's test with Bonferroni correction for multiple tests.

Then, we investigated whether the studied variables, i.e., age and sex, impacted the occurrence of each self-reported strategy. Therefore, two multinomial log-linear models were assessed (one model for words and one model for pseudowords) with the four self-reported strategies as the multiclass outcome and the two variables of interest as the predictors (*nnet::multinom* function, Venables and Ripley, 2002).

Modeling Memory Performance by Self-Reported Strategies

To estimate the effects of the self-reported strategies on word and pseudoword memory performance at each memory stage, we implemented a linear mixed model. A 3-way interaction of strategy (4 levels: Listening, Position, Meaning, and Dual) * type of item (2 levels: words and pseudowords) * memory stage (3 levels: STR, LOT and DR) was introduced in the model as a fixed effect. The random effects of type of item and memory stage were fitted on the intercept at the subject level. A stepwise backward strategy for model selection using the AIC was applied to the linear mixed model described above to determine the best-fitted model. The ANOVA components were calculated using the Kenward–Roger approximation. The residuals were visually inspected to assess normality and homoscedasticity. The proportion of variance explained by the fixed effects and their interactions was estimated with the marginal R^2. All pairwise *post hoc* comparisons were corrected for multiple comparisons using an FDR-controlling procedure.

Causal Mediation Analysis (CMA)

Causal mediation analysis was conducted to investigate the contribution of the self-reported strategies to the relationships between age or sex and memory scores. Causal mediation analysis decomposes the total effect of a predictor into direct (i.e., the effect of the predictor on the dependent variable adjusted for the predictor–mediator relationship) and indirect (i.e., the mediator effect) effects. The MBESS library was implemented (*MBESS::mediation*, Kelley, 2019) with a bias-corrected bootstrapped 95% CI with 10,000 samples. The effect size was estimated by kappa squared (κ^2), which corresponds to the proportion of the maximum possible indirect effect that could have occurred (Preacher and Kelley, 2011). Since the first method of calculation was criticized (Wen and Fan, 2015), here, κ^2 was calculated according to Talloen's proposition (Talloen W. Effect size measures for mediation models: a critical evaluation of κ2. Unpublished results, 2015) and implemented in R with the *MaxIE* function provided by the author.

RESULTS

Characterization of Words Memory Scores and Comparison With Pseudoword Memory Scores

First, we characterized and compared the memory scores of the words with those of the pseudowords. The first result was the number of recalled words that was significantly higher than the number of pseudowords at each memory stage, corresponding to a generalized semantic benefit (**Table 2**).

Effects of Age and Sex on Memory Performance

After adjusting for education, vocabulary scope, rhyming, WMC, and handedness, the effects of age and sex on the three memory stages of the words and pseudowords were modeled with a linear mixed model. The final model, which was obtained by a stepwise backward strategy, explained 27% of the variance, and 13.3% of the variance was due to fixed-effect factors.

TABLE 2 | Descriptive data of the studied sample and memory measures.

Studied sample (N = 447)	Mean (SD)/% (n)	[Range]
Age (years)	26.64 (7.60)	[18.09–57.22]
Sex	W: 51 % (228)	
	M: 49 % (219)	
Confounding factors		
Handedness	Left: 45.41 % (203)	
	Right: 54.59 % (244)	
Education (years)	15.27 (2.50)	[8–20]
Vocabulary scope (max. 44)	27.47 (4.25)	[15–36]
Rhyming (max. 80)	67.02 (6.08)	[44–77]
WMC (max. 6)	4.17 (0.99)	[2–6]
Verbal list test (RAVLT)		
STR (trial 1, max. 18)	7.99 (2.00)*	[2–16]
LOT (trials 2 + 3)−(trial 1 × 2)	9.72 (3.21)*	[1–18]
DR (% of max. trial)	90.49 (10.80)*	[50–121.43]
Pseudoword list test		
STR (trial 1, max. 15)	2.91 (1.40)	[0–9]
LOT (trials 2 + 3)−(trial 1 × 2)	6.97 (3.78)	[-2 to 17]
DR (% of max. trial)	62.30 (24.73)	[0 - 118.18]

*Continuous variables are summarized as the mean ± SD, and categorical variables are summarized as % (n); WMC, working memory capacity; STR, short-term recall; LOT, learning over trials; DR, delayed recall; *significant differences between word and pseudoword scores, as assessed by a paired samples Bonferroni t-test corrected for the multiplicity of tests (p < 0.016). nota bene: The same significant differences were observed when STR and LOT were expressed as the percentage of list length.*

Age Effect on Pseudoword and Word Memory According to the Memory Stage

A 3-way interaction among age, type of item and memory stage was observed ($F_{(2,1332)}$ = 3.61; p = 0.03; **Figure 1** and **Table 3**). A general decline in performance with age was visible in the 3 pseudoword memory stages but was limited to STR and DR for word memory. Thus, with increasing age, word meaning improved word memory scores relative to pseudoword memory scores for LOT (ß = 0.249, $t_{(1,1332)}$ = 4.44, p < 0.0001) and, to a lesser extent, for DR (ß = 0.119, $t_{(1,1332)}$ = 2.07, p = 0.04).

Sex Effect on Pseudoword and Word Memory According to the Memory Stage

A 3-way interaction among sex, type of item, and memory stage was observed ($F_{(2,1332)}$ = 9.12; p < 0.0001; **Figure 2** and **Table 4**). The word memory scores were significantly higher than the pseudoword memory scores exclusively among women for STR (FDR-corrected *post hoc* test: words: $t_{(1,1332)}$ = 2.51, p = 0.012). In contrast, the STR of pseudowords was higher than the STR of words among men (FDR-corrected *post hoc* test: words: $t_{(1,1332)}$ = −2.57, p = 0.010). From the perspective of sex contrasts, women performed significantly better than men on word STR and DR (FDR-corrected *post hoc* tests: STR: $t_{(1,2541)}$ = 3.95, p < 0.0001; DR: $t_{(1,2541)}$ = 1.97, p = 0.049) and on pseudoword LOT and DR (FDR-corrected *post hoc* tests: LOT: $t_{(1,2541)}$ = 3.92, p < 0.0001; DR: $t_{(1,2541)}$ = 2.93, p = 0.003).

Characterization of the Self-Reported Strategies

The distribution of the four self-reported strategies for the memory of words (N = 444) and pseudowords (N = 442) is shown in **Table 5**.

Strikingly, the strategies used to retain the list of pseudowords were primarily Meaning based, as follows: 38.4% of the participants used the Meaning strategy, and 19.3% of the participants used this strategy in association with the Position strategy. The Position strategy alone was used by 18% of the participants. For the RAVLT, the dominant strategy was Meaning, as 32.3% of the participants used this strategy alone, 32.5% of the participants used this strategy in association with the Position strategy, and 21.1% of the participants exclusively used the Position strategy. The distribution of the pseudoword strategies significantly differed from the distribution of word strategies (McNemar's $\chi^2_{(6)}$ = 24.40, p < 0.0001). Compared to the words, the pseudowords elicited a greater use of the Listening strategy (McNemar's $\chi^2_{(1)}$ = 27.13, p < 0.0001) and less use of the Dual strategy (McNemar's $\chi^2_{(1)}$ = 55.65, p < 0.0001).

Effects of the Self-Reported Strategies on Memory Performance

The effects of the self-reported strategies on the 3 stages of word and pseudoword memory were modeled in a linear mixed model. The final model, which was obtained by a stepwise backward strategy, explained 28% of the variance, and 5.6% of the variance was attributable to fixed-effect factors.

The interaction between the self-reported strategies and the type of item had a major significant effect ($F_{(3,662)}$ = 4.87; p = 0.002), which is clearly visible in **Figure 3** and **Table 6**. Indeed, the Meaning and Dual strategies both improved word memory performance compared to the Listening strategy (respective FDR-corrected *post hoc* tests: $t_{(1,401)}$ = 3.17, p = 0.003 and $t_{(1,401)}$ = 3.42, p = 0.003) and the Position strategy (respective FDR-corrected *post hoc* tests: $t_{(1,401)}$ = 2.94, p = 0.005 and $t_{(1,401)}$ = 3.21, p = 0.003). Concerning the pseudowords, all 3 advanced strategies improved memory performance relative to the Listening strategy (FDR-corrected *post hoc* tests: Position: $t_{(1,401)}$ = 3.11, p = 0.003, Meaning: $t_{(1,401)}$ = 3.58, p = 0.001, Dual: $t_{(1,401)}$ = 6.16, p < 0.0001). Moreover, regarding the pseudowords, use of the Dual strategy significantly increased the memory score compared to the Position strategy (FDR-corrected *post hoc* tests: $t_{(1,401)}$ = 2.87, p = 0.005) and the Meaning strategy (FDR-corrected *post hoc* tests: $t_{(1,401)}$ = 3.40, p = 0.001), suggesting that a potentiating effect occurs when the Position and Meaning strategies are combined.

Mediation of the Age and Sex Effect on Memory Performance by Self-Reported Strategies

First, we assessed the impact of age and sex on the likelihood of using the 4 self-reported strategies compared to each other. The main outcomes are summarized in **Table 7**. Relative to use of the Listening strategy, the likelihood of using any of the three advanced strategies to retain both the words and
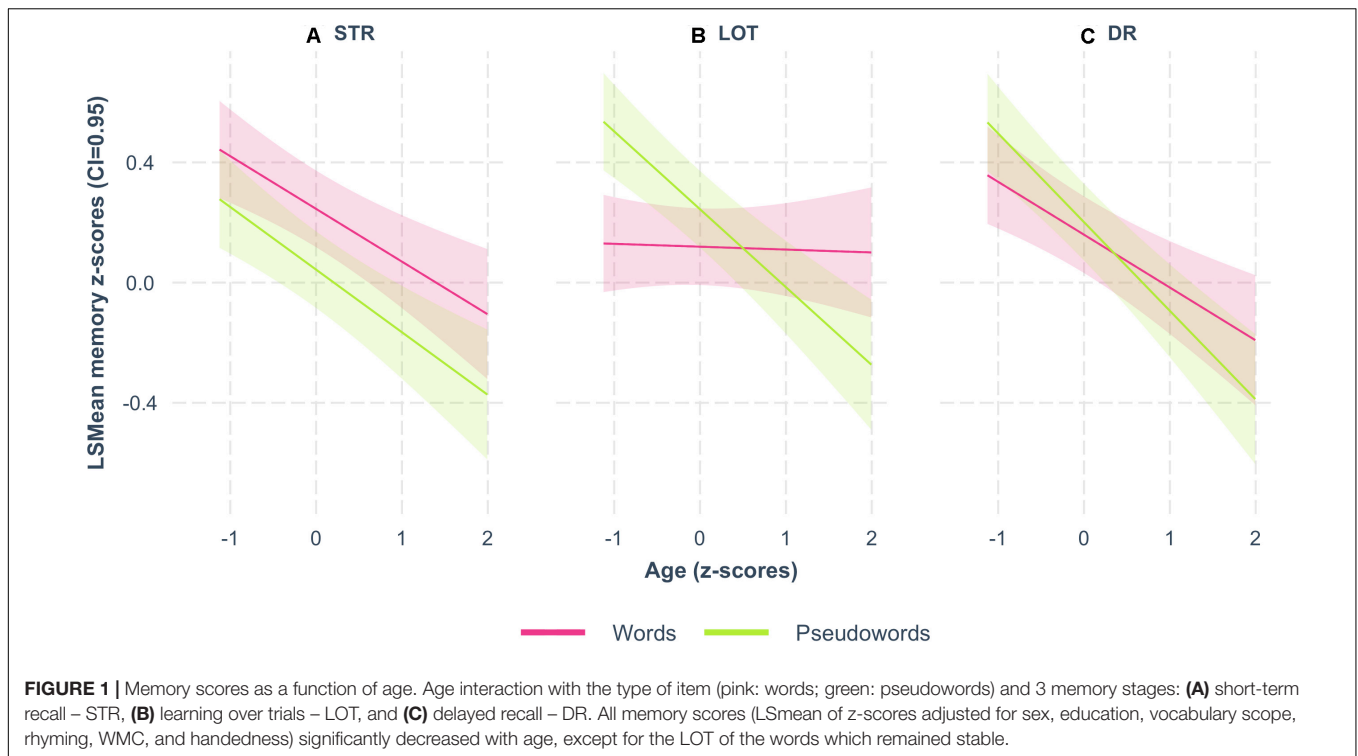
FIGURE 1 | Memory scores as a function of age. Age interaction with the type of item (pink: words; green: pseudowords) and 3 memory stages: **(A)** short-term recall – STR, **(B)** learning over trials – LOT, and **(C)** delayed recall – DR. All memory scores (LSmean of z-scores adjusted for sex, education, vocabulary scope, rhyming, WMC, and handedness) significantly decreased with age, except for the LOT of the words which remained stable.

TABLE 3 | Memory performance as a function of age for words and pseudowords at each memory stage after adjusting for sex, education, vocabulary scope, rhyming, WMC, and handedness.

	STR: short-term recall				LOT: learning over trials				DR: delayed recall			
(A) Age (+1SD)	*ß*	*SE*	$t_{(2508)}$	*p*	*ß*	*SE*	$t_{(2508)}$	*p*	*ß*	*SE*	$t_{(2508)}$	*p*
Words	−0.175	0.04	−3.92	0.0001	−0.010	0.04	−0.21	0.831	−0.176	0.04	−3.93	0.0001
Pseudowords	−0.208	0.04	−4.65	<0.0001	−0.259	0.04	−5.78	<0.0001	−0.295	0.04	−6.58	<0.0001
(B) Words > Pseudowords	*ß*	*SE*	$t_{(1332)}$	*p*	*ß*	*SE*	$t_{(1332)}$	*p*	*ß*	*SE*	$t_{(1332)}$	*p*
	0.033	0.05	0.57	0.567	0.249	0.05	4.44	<0.0001	0.119	0.05	2.07	0.038

(A) FDR-corrected post hoc estimation of the word and pseudoword slopes associated to 1 SD of age. (B) The FDR-corrected post hoc pairwise comparisons of the word and pseudoword slopes revealed a significant difference for LOT and DR.

pseudowords decreased with age. Moreover, comparisons to the Position strategy revealed an influence of sex. Regarding words, being a woman increased the likelihood of using the Meaning and the Dual strategies. Regarding pseudowords, being a woman increased the likelihood of using the Meaning strategy.

Then, CMA was conducted to assess whether the self-reported strategies mediated the age and sex effects on memory performance (**Table 8**). Regarding words, the following findings were observed: (1) the sex effect on STR (higher scores among women) was mediated by women's greater use of the Meaning strategy (used alone or in combination with the Position strategy) and (2) the age effect on STR and DR (decreasing scores with age) was mediated by the reduced use of an advanced strategy (i.e., all strategies but Listening) in older participants. Similarly, the decrease of all pseudoword memory scores with age was mediated by the reduced use of an advanced strategy (i.e., all strategies but Listening). Notably, the effect sizes, as measured by κ^2, varied between 2.4 and 8.1% of the maximum possible indirect effect and can be considered "small."

DISCUSSION

As might be expected, compared to learning pseudowords, learning words (known meaningful items) resulted in improved performance at each memory stage, from STR to DR, via LOT (**Table 2**). Since in supraspan lists, more items are *a priori* more difficult to memorize, the longest length of the list cannot explain the better memory scores for words. This word advantage could arise from a detrimental effect of pseudowords' novelty, which would have an especially strong impact on STR. However, as the word advantage decreased in the following trials, as shown by the proximity between word and pseudoword LOT, the pseudowords became more familiar
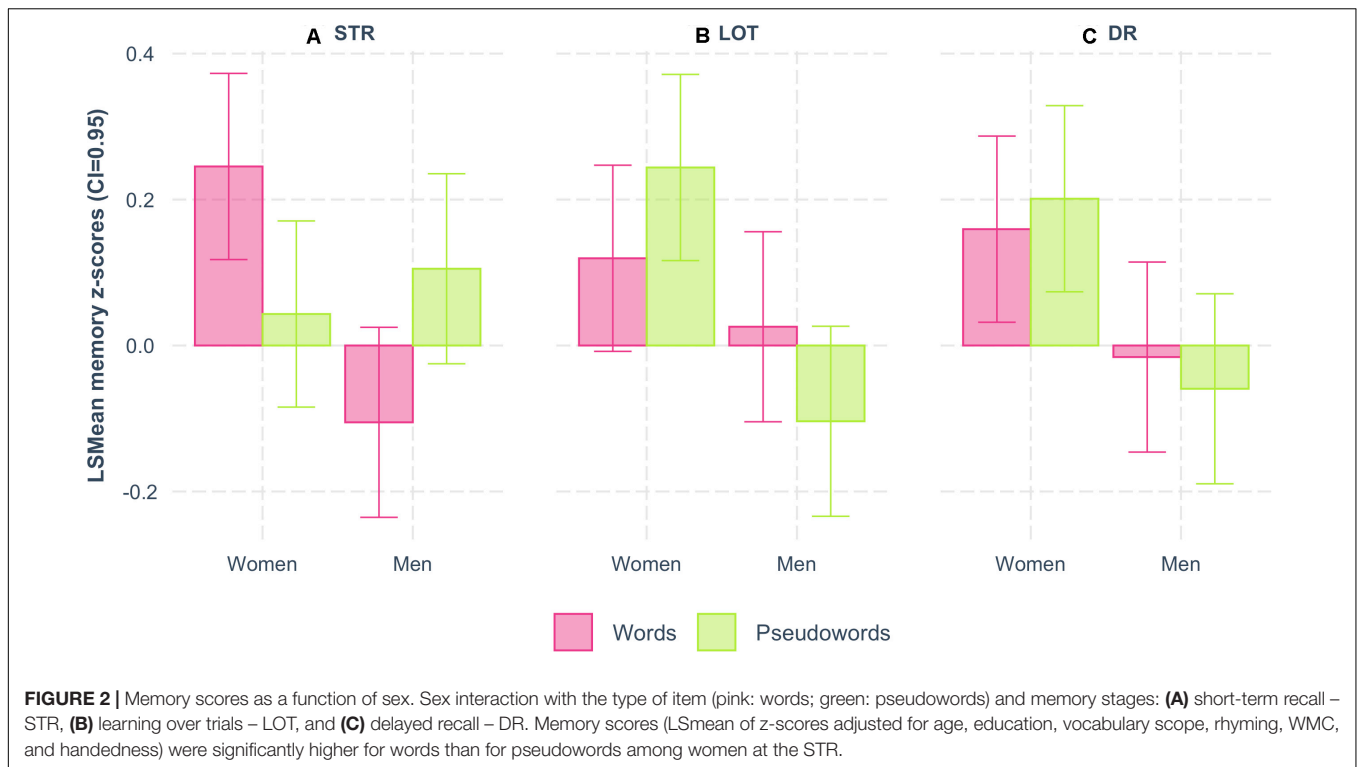
FIGURE 2 | Memory scores as a function of sex. Sex interaction with the type of item (pink: words; green: pseudowords) and memory stages: **(A)** short-term recall – STR, **(B)** learning over trials – LOT, and **(C)** delayed recall – DR. Memory scores (LSmean of z-scores adjusted for age, education, vocabulary scope, rhyming, WMC, and handedness) were significantly higher for words than for pseudowords among women at the STR.

TABLE 4 | Memory performance as a function of sex and its interaction with the type of item and memory stage after adjusting for age, education, vocabulary scope, rhyming, WMC, and handedness.

Sex	STR: short-term recall				LOT: learning over trials				DR: delayed recall			
(A) Words > Pseudowords	**ß**	**SE**	**$t_{(1332)}$**	**p**	**ß**	**SE**	**$t_{(1332)}$**	**p**	**ß**	**SE**	**$t_{(1332)}$**	**p**
Women	0.202	0.08	2.51	0.012	−0.124	0.08	−1.55	0.121	−0.042	0.08	−0.52	0.603
Men	−0.210	0.08	−2.57	0.010	0.129	0.08	1.58	0.114	0.043	0.08	0.53	0.596
(B) Women > Men	**ß**	**SE**	**$t_{(2541)}$**	**p**	**ß**	**SE**	**$t_{(2541)}$**	**p**	**ß**	**SE**	**$t_{(2541)}$**	**p**
Words	0.35	0.06	3.95	<0.0001	0.094	0.06	1.06	0.291	0.175	0.06	1.97	0.049
Pseudowords	−0.062	0.06	−0.69	0.485	0.348	0.06	3.92	<0.0001	0.260	0.06	2.93	0.003

(A) FDR-corrected pairwise comparisons of word and pseudoword memory performance among women and men at each memory stage. (B) FDR-corrected pairwise comparisons of women's and men's memory performance for words and pseudowords at each memory stage.

with repetition. The word advantage observed in the latest stage of memorization thus indicates that words benefit from an effect other than mere familiarity, one that is probably related to the meaning they carry.

TABLE 5 | Self-reported strategy distribution according to the type of item.

Strategy type	Pseudowords (%)	Words (%)	McNemar's χ^2	p
Listening	24.3	14.1	24.40	<0.0001
Position	18.0	21.1	2.23	0.5424
Meaning	38.4	32.3	5.17	0.0919
Dual (Position + Meaning)	19.3	32.5	27.13	<0.0001

The differences in the proportions of strategy use between pseudowords and words were tested with McNemar's Bonferroni test corrected for four tests.

Our observation of the age effects on word memory performance (**Table 3** and **Figure 1**) reproduced numerous previous findings showing that age (16–86 years) was associated with a decrease in both the STR and DR of words but did not have a visible effect on LOT (Geffen et al., 1990; Mitrushina et al., 1991; Poreh, 2005; Van Der Elst et al., 2005). In addition, our results showed that age systematically led to a decrease in pseudoword memory performance for each memory stage. The dissimilar effects of age on word and pseudoword performance were particularly important for LOT suggesting that the word-related gain lasts longer, specifically during the learning process.

We also examined the sex effect since the literature has regularly reported for many years that women perform better than men on STR and DR (Bolla-Wilson and Bleecker, 1986; Geffen et al., 1990; Aartsen et al., 2004; Van Der Elst et al., 2005; Gale et al., 2007; Messinis et al., 2007; Badcock et al., 2011). From
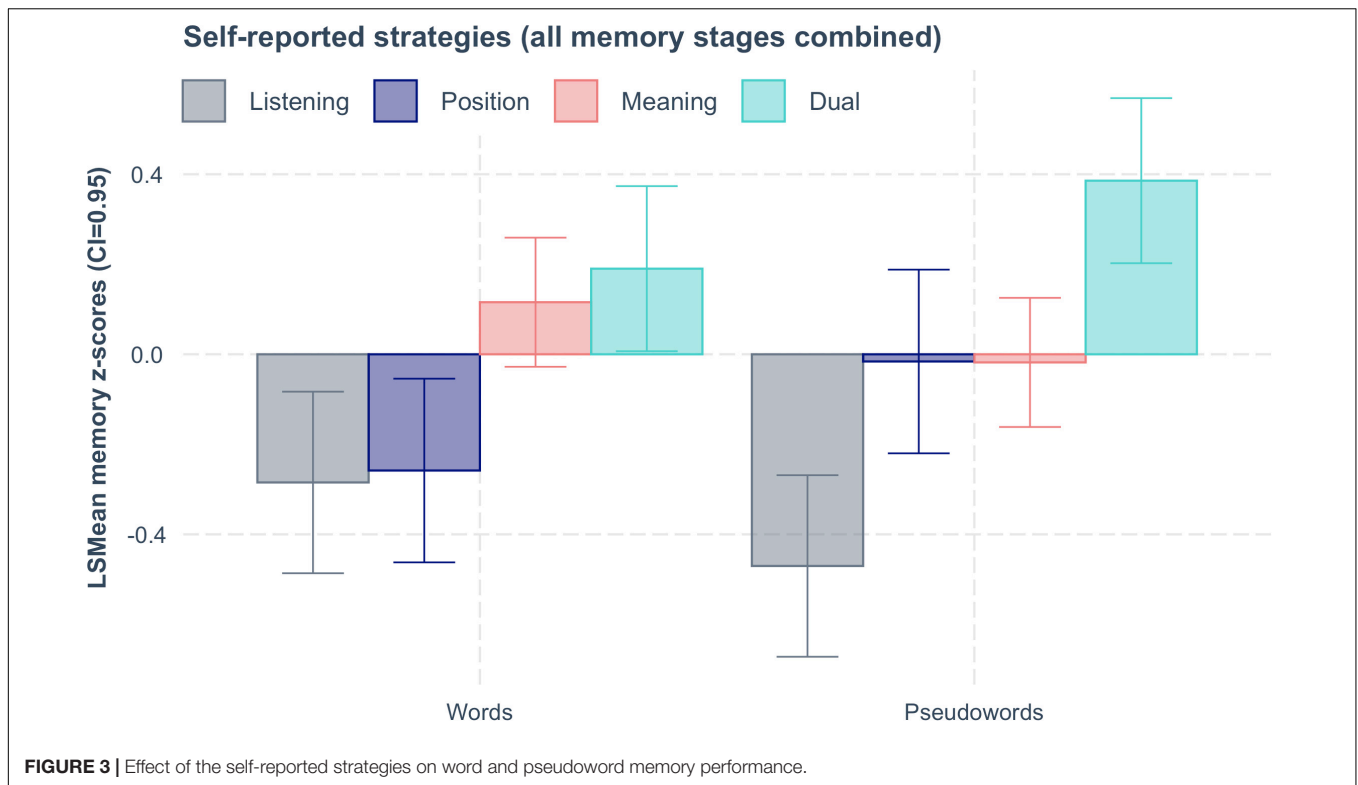
FIGURE 3 | Effect of the self-reported strategies on word and pseudoword memory performance.

TABLE 6 | Memory performance as a function of self-reported strategy and its interaction with the type of item.

Strategies	Position				Meaning				Dual			
	β	SE	$t_{(401)}$	p	β	SE	$t_{(401)}$	p	β	SE	$t_{(401)}$	p
(A) Listening <												
Words	0.026	0.15	0.18	0.856	0.400	0.13	3.17	0.003	0.475	0.14	3.42	0.003
Pseudowords	0.454	0.15	3.11	0.003	0.452	0.13	3.58	0.001	0.856	0.14	6.16	<0.0001
(B) Position <												
Words					0.374	0.13	2.94	0.005	0.448	0.14	3.21	0.003
Pseudowords					−0.002	0.13	−0.02	0.987	0.402	0.14	2.87	0.005
(C) Meaning <												
Words									0.074	0.12	0.63	0.636
Pseudowords									0.404	0.12	3.40	0.001

FDR-corrected pairwise comparisons of memory performance for each strategy for words and pseudowords: (A) Listening as ref., (B) Position as ref., (C) Meaning as ref.

the perspective of this sex contrast (**Table 4** and **Figure 2**), our results reproduced these previous findings: women performed better than men on the STR and DR of words, but there was no visible sex difference in word LOT performance (Teruya et al., 2009). Regarding the difference in performance for words and pseudowords, the gain provided by the words emerged only among the women on the STR (**Figure 2**). This result demonstrates that the word gain can appear during early memory stages, which supplements recent studies highlighting the nature and importance of short-term semantic memory (Campoy et al., 2015; Aizpurua and Koutstaal, 2018).

To better understand the explicit mnemonics involved in word and pseudoword memory, we determined the strategies the participants used to retain the lists of words and pseudowords.

We found that words and pseudowords elicited the same types of strategies (**Table 1**). Two main strategies were reported, namely, a strategy based on associations of meanings (lexical/semantic binding) of the items and a strategy based on the temporal contiguity (position on the list) of the items. These two strategies, called Meaning and Position in the present study, correspond to memory strategies previously described in free recall studies as *subjective* and *serial clustering* (Meijs et al., 2013) or *semantic proximity effects* and *temporal contiguity* (Sederberg et al., 2010). These two strategies were contrasted with a third strategy, namely, the Listening strategy, which has also been previously described in free recall tasks as a *rehearsal* mnemonic technique (Herrmann, 1987). These two strategies were also used in combination, an approach that we called the Dual strategy. Thus,

TABLE 7 | Adjusted odds ratio (aOR) and 95% confidence interval (CI) derived from the multinomial logistic model comparing each strategy to the Listening and Position strategies.

	Words				**Pseudowords**		
	LISTENING as ref.	**aOR**	**95% CI**		**LISTENING as ref.**	**aOR**	**95% CI**
Age	**Position**	**0.95**	**0.91–0.98**	**Age**	**Position**	**0.95**	**0.92–0.99**
	Meaning	**0.96**	**0.92–0.99**		**Meaning**	**0.95**	**0.92–0.98**
	Dual	**0.93**	**0.90–0.97**		**Dual**	**0.92**	**0.88–0.97**
Women	Position	0.83	0.43–1.62	**Women**	Position	0.83	0.46–1.52
	Meaning	1.52	0.82–2.81		Meaning	1.46	0.88–2.42
	Dual	1.7	0.91–3.18		Dual	1.31	0.72–2.38
	POSITION as ref.	**aOR**	**95% CI**		**POSITION as ref.**	**aOR**	**95% CI**
Age	**Listening**	**1.06**	**1.01–1.10**	**Age**	**Listening**	**1.05**	**1.01–1.09**
	Meaning	1.01	0.97–1.05		Meaning	0.99	0.96–1.04
	Dual	0.98	0.95–1.03		Dual	0.96	0.91–1.02
Women	Listening	1.2	0.62–2.34	**Women**	Listening	1.2	0.66–2.18
	Meaning	**1.83**	**1.06–3.13**		**Meaning**	**1.74**	**1.01–3.00**
	Dual	**2.05**	**1.19–3.52**		Dual	1.56	0.84–2.92

When the CI excludes the value 1 (in bold), the null hypothesis is rejected, indicating a significantly lower (<1) or higher (>1) likelihood of occurrence compared to the reference level.

from the complexity perspective, the four strategies could be ordered as follows: the Listening strategy was the simplest, a minimal strategy induced by the item presentation; the Position strategy was equivalent to the Meaning strategy; and the Dual strategy was the most complex strategy, combining two strategies.

Regarding the distribution of the four strategies according to type of item (**Table 5**), word memory relied on the use of more complex strategies than pseudoword memory, which led to a reduced use of the Listening strategy and an increased use of the Dual strategy. The Meaning strategy alone or in combination was the dominant strategy used for word memory (64% of the respondents) and, quite surprisingly, pseudoword memory (57% of the respondents). The main difference between the words and pseudowords concerned the nature of the Meaning strategy, which consisted of thematic and taxonomic associations for the words and lexical associations for the pseudowords. This result suggests that auditory supraspan list-learning tests encourage memory based on associations with familiar and meaningful knowledge in a hierarchical manner (depending on the cues the item provides), from phonemes to lexicon for pseudowords and from lexicon to semantic binding for words. This interpretation is consistent with a previous finding that participants performed equally well on the RAVLT and CVLT when the semantic categorization of words was experimentally forced in the CVLT (Crossen and Wiens, 1994). We cannot excluded the possibility that the fixed order of list presentation (words followed by pseudowords ~3 hours later) may have favored the use of a Meaning strategy for pseudowords and constitutes a potential limitation to this study. Nevertheless, the subtypes of the Meaning strategy were not strictly the same for the words and pseudowords (**Table 1**). Thus, it appears that exposure to known phonological entities, regardless of their meaning, encouraged the use of lexico-semantic associations.

Subsequently, we found that the effectiveness of the self-reported strategies varied between word and pseudoword memory tasks (**Table 6** and **Figure 3**). The Listening strategy led to the lowest word and pseudoword memory performance. However, the Meaning strategy, used alone or in combination with the Position strategy, improved word memory, while the Dual strategy was the most efficient for the pseudoword memory, suggesting that for pseudowords, the Meaning and Position strategies potentiated each other. According to Worthen and Hunt, list-learning mnemonic techniques can be divided into *organization* mnemonic processes, which emphasize intralist associations (i.e., all information to be remembered is linked together), and *elaboration* mnemonic processes, which rely on extralist cues (e.g., meaningfulness, mental imagery) (Worthen and Hunt, 2017). Furthermore, these authors argued that the combined use of the organization and elaboration strategies is the most effective mnemonic. Here, the Position strategy was analogous to an organization mnemonic, while the Meaning strategy was analogous to an elaboration mnemonic. The combination of organization and elaboration mnemonics appeared to be the most effective strategy for pseudoword memory (**Figure 3** right). However, this did not apply to word memory, as only the elaboration mnemonic enabled the best performance (**Figure 3** left). Moreover, concerning word memory, the Position strategy led to performances as low as those resulting from the use of Listening strategy, which is congruent with the previous findings showing that rote rehearsal is the least effective method for free recall tasks (Herrmann, 1987). Taken together, these two observations strongly suggest that the Meaning strategy was an extremely efficient mnemonic that had the potential to maximize verbal list-learning memory performance.

Finally, we were interested in determining whether the age and sex effects on memory performance could have been

TABLE 8 | Summary of the significant CMA (causal mediation analyses) of the standardized memory scores at each memory stage; STR, short-term recall; LOT, learning over trials; DR, delayed recall; maxIE, maximum possible indirect effect.

	Effect	ß	95% CI	maxIE	κ^2
Words					
Women → Meaning strategy + Dual strategy → **STR**	Indirect	0.0603	0.0189 to 0.1157	0.74	0.0815
	Direct	0.2599	0.0785 to 0.4412		
Age → Advanced strategies → **STR**	Indirect	−0.0296	−0.0644 to −0.0098	−0.8000	0.037
	Direct	−0.192	−0.2845 to −0.0995		
Age → Advanced strategies → **DR**	Indirect	−0.0191	−0.0426 to −0.0013	−0.8000	0.0239
	Direct	−0.1993	−0.2924 to −0.1062		
Pseudowords					
Age → Advanced strategies → **STR**	Indirect	−0.0365	−0.0685 to −0.0154	−0.7821	0.0467
	Direct	−0.2106	−0.3029 to −0.1184		
Age → Advanced strategies → **LOT**	Indirect	−0.0455	−0.0817 to −0.0212	−0.7400	0.0615
	Direct	−0.2596	−0.3495 to −0.1696		
Age → Advanced strategies → **DR**	Indirect	−0.0472	−0.0849 to −0.0227	−0.7100	0.0665
	Direct	−0.2894	−0.3781 to −0.2006		

mediated by the previous self-reported strategies. We found that aging significantly decreased the likelihood of using any of the advanced strategies, i.e., all strategies other than Listening, for word and pseudoword memory (**Table 7**), which, in turn, impaired memory performance, as assessed by the CMAs (**Table 8**). This result strongly suggests that an increase in age produces a nonspecific decrease in the use of advanced strategies. We also observed that being a woman increased the likelihood of using the Meaning and Dual strategies compared to the Position strategy, for word memory (**Table 7**) and that this effect could to some extent mediate women's superior word STR performance (**Table 8**). Thus, the gain in STR performance provided by the words could be specifically related to the fact that women preferentially use the Meaning strategy (alone or in combination with the Position strategy). This result is congruent with previous observations obtained using the CVLT (Kramer et al., 1988, 1997) and supports Andreano and Cahill's hypothesis that women's verbal memory advantage depends on encoding at the semantic level (Andreano and Cahill, 2009).

From a neuropsychological perspective, the distinction between normal and pathological aging and the early detection and identification of neurodegenerative diseases are core features of investigations. Since we found that RAVLT scores can capture semantic strategies, we can assume that the RAVLT enables the detection of the early semantic impairment, particularly word-finding difficulties, that characterizes AD (Herlitz and Viitanen, 1991; Croisile et al., 1996; de Lira et al., 2011; Verma and Howard, 2012). Extending this idea, the semantic sensitivity of the RAVLT might explain its previously observed ability to detect the disease in its early stages and, notably, to discriminate AD from MCI (Marra et al., 2000; Estévez-González et al., 2003; Goryawala et al., 2015; Bauer et al., 2018). Moreover, we observed that the word memory advantage of women was mediated by the use of the Meaning strategy (**Table 8**). This sex-specific mediation can lead to a sex-specific cognitive advantage that compensates for the neurological evolution of AD and, as previously hypothesized by Sundermann et al. (2017), interferes with the early detection

of the disease in women. Therefore, when the RAVLT is used for diagnostic purposes, we join the authors in advocating the establishment of sex-specific cutoff scores (Nebel et al., 2018) or sex-adapted performance gaps between healthy individuals and those with dementia (Chapman et al., 2011; Gale et al., 2016).

Since the introduction of the RAVLT in the 1950s, new theories about memory have emerged. One of the most widely recognized is Tulving's theory, which distinguishes nondeclarative (implicit) memory from declarative (explicit) memory. Moreover, Tulving describes two types of declarative memory, as follows: (1) *episodic memory*, corresponding to personal memories (or contextualized events) and referring to recollection, a key concept related to autonoetic consciousness (a subjective sense of self over time) and (2) *semantic memory*, corresponding to general knowledge and referring to familiarity, which is associated with noetic consciousness (Tulving, 2002). At this stage, it could be useful to relate Tulving's model of declarative memory to our observations and to discuss the ability of the RAVLT to assess Tulving's memory systems in light of our results.

Tulving previously specified that "*there is no necessary correlation between behavior and conscious experience and in that sense the traditional research* (i.e*., such sterile situations as list-learning experiments) was not concerned with episodic memory*" (Tulving, 2002). Nevertheless, in the early and mid-1990s, free recall tasks (used as declarative memory tests) and RAVLT scores (DR in particular) began to be associated with the assessment of episodic memory (Roediger, 1990; Litvan et al., 1991; Moffot et al., 1994; Eustache et al., 1995; Kirkby et al., 1995), and this association became increasingly frequent over time, occurring in 13–20% (PubMed) or 33–43% (ScienceDirect) of the RAVLT literature over the last 6 years (keyword cross-search from January 2020). Currently, it is not unusual to find a direct association between the RAVLT and the assessment of verbal episodic memory (Moradi et al., 2017; Barulli et al., 2019; Putcha et al., 2019; Sudo et al., 2019) or even to find the RAVLT being used to validate other verbal episodic memory

tests (Morrison et al., 2018). Since episodic memory implies explicit retrieval (Tulving, 1995), the exploration of self-reported strategies is a direct way to investigate whether episodic memory is assessed by the RAVLT. Indeed, a strategy linked to episodic memory could make some autonoetic references to the context of the test (e.g., a participant associating "handle" with the "handle of the door *in the experiment room*") or an item's association with a personal contextualized recollection (e.g., a participant associating "fireplace" with the "fireplace *from his childhood*"). This type of autonoetic strategy did not emerge as a regularly used strategy in our results, leading to the conclusion that the RAVLT seems unable to assess episodic memory to a reasonable degree, which is similar to the conclusion reached by previous authors (Tulving, 1985, 2002; Van der Linden, 2004; Desgranges and Eustache, 2011). Consistent with Gavett and Horwitz's assertion that list-learning test interpretation lacks the construct validity to allowed unbiased estimates of episodic memory ability (Gavett and Horwitz, 2012), our results supplement those reported by Casaletto et al. (2017), who focused on total immediate recall performance during learning trials and concluded that interpreting learning scores as equivalent to episodic memory may be erroneous. To avoid any confusion regarding the purpose of the RAVLT, and considering its widespread use, we encourage the use of simple terminology, i.e., "(supraspan) verbal memory," to briefly characterize the RAVLT assessment. In addition, it seems appropriate to extend this caution to all supraspan verbal list-learning tests.

Within the framework of Tulving's memory model, the present results lead to consider the RAVLT scores as being affected by semantic memory. Indeed, the present results show that general knowledge of the meaning of words (the thematic and taxonomic associations used in the Meaning strategy) maximize performance when used as a mnemonic (**Figure 3**). Moreover, the Meaning strategy was widely used among the participants (**Table 5**) and was found to mediate women's advantage regarding the STR of words (**Table 8**). Therefore, RAVLT performance appears to be highly dependent on this general knowledge and, thus, on semantic memory. Nonetheless, the self-reported strategies explained only 28% (5.6% imputable to the fixed effects) of the variance in RAVLT performance and mediated 2–8% of the age and sex effects on memory performance. These effect sizes suggest that some memory processes escaped the participants' self-assessment and that the RAVLT scores surely assessed some implicit mechanisms. This hypothesis is compliant with Tulving's definition of semantic memory, which is declarative knowledge characterized by implicit retrieval (Tulving, 1995). In addition, a quick review of the literature reveals the possible involvement of early implicit encoding processes, such as working memory and executive functions. Although the age and sex effects on memory performance that we observed were adjusted for WMC, other dimensions of working memory and executive functions may be involved. A recent study demonstrated that semantic encoding is automatic in verbal short-term memory (Campoy et al., 2015). Furthermore, the recency effect, which plays a significant role in RAVLT scores, is largely driven by verbal attention (Griffin et al., 2017). The recency effect has been conceptualized in immediate free recall as implicit learning coupled with a particular mode of retrieval that may, but need not, be conscious and explicit (Baddeley and Hitch, 1993). Finally, the present observation that aging nonspecifically decreases the use of advanced self-reported strategies (**Table 7**) also argues for the involvement of executive functions. This interpretation is consistent with previous observations that normal aging is associated with more difficulty accessing lexico-semantic operations and representations due to a slowdown in executive functions rather than concept loss per se (Baciu et al., 2016). A recent study also showed selective inability to recall RAVLT midlist items in patients with a selective mild executive function deficit (Consonni et al., 2017). Moreover, some authors noted the probable combined effect of attentional and short-term memory processes on the immediate recall score (Gavett and Horwitz, 2012) and a high degree of overlap between verbal memory and executive functioning (Duff et al., 2005). While further studies are needed to clarify the nature of these associations, our results contribute to previous studies suggesting the existence of an association between RAVLT scores and working memory or executive functions.

In conclusion, our results showed that word meaning provides a significant gain in recall performance on the RAVLT. In particular, word meaning enables access to meaning-based memory strategies that optimize verbal memory performance. Moreover, the beneficial role of the Position memory strategy for pseudoword recall was potentiated by a lexical-based memory strategy. These two observations strongly suggest that pseudoword and word-list learning memory performance depends on strategies that are based on general semantic knowledge. Within the framework of Tulving's declarative memory model, our results indicate that the verbal list-learning test free recall scores are affected by semantic memory. Since the self-reported strategies were not autonoetic, we conclude that the RAVLT is not suitable for episodic memory assessment as it is increasingly referred to in the literature.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comité de Protection des Personnes Nord-Ouest, France. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS

This manuscript has been greatly benefited from the precious support and thoughtful comments of Nathalie Tzourio-Mazoyer, Bernard Mazoyer, Michel Thiebault de Schotten, and Laurent Petit.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.02043/full#supplementary-material

REFERENCES

Aartsen, M. J., Martin, M., and Zimprich, D. (2004). Gender differences in level and change in cognitive functioning. *Gerontology* 50, 35–38. doi: 10.1159/000074387

Aizpurua, A., and Koutstaal, W. (2018). A new index of semantic short-term memory: development and validation of the conceptual span task in Spanish. *PLoS ONE* 13:e0209368. doi: 10.1371/journal.pone.0209368

Andreano, J. M., and Cahill, L. (2009). Sex influences on the neurobiology of learning and memory. *Learn. Mem.* 16, 248–266. doi: 10.1101/lm.918309

Baciu, M., Boudiaf, N., Cousin, E., Perrone-Bertolotti, M., Pichat, C., Fournet, N., et al. (2016). Functional MRI evidence for the decline of word retrieval and generation during normal aging. *Age* 38:3. doi: 10.1007/s11357-015-9857-y

Badcock, J. C., Dragovic, M., Dawson, L., and Jones, R. (2011). Normative data for Rey's auditory verbal learning test in individuals with schizophrenia. *Arch. Clin. Neuropsychol.* 26, 205–213. doi: 10.1093/arclin/acr005

Baddeley, A. D., and Hitch, G. (1993). The recency effect: implicit learning with explicit retrieval? *Mem. Cognit.* 21, 146–155. doi: 10.3758/BF03202726

Barulli, M. R., Piccininni, M., Di Dio, C., Musarò, C., Grasso, A., Tursi, M., et al. (2019). Episodic memory and learning rates in amyotrophic lateral sclerosis without dementia. *Cortex* 117, 257–265. doi: 10.1016/j.cortex.2019.03.003

Bauer, C., Cabral, H., and Killiany, R. (2018). Multimodal discrimination between normal aging, mild cognitive impairment and Alzheimer's disease and prediction of cognitive decline. *Diagnostics* 8:14. doi: 10.3390/diagnostics8010014

Binois, D. R., and Pichot, D. P. (1956). *Test de Vocabulaire: Manuel d'Application*. Paris: Éditions du Centre de Psychologie Appliquée.

Bleecker, M. L., Bolla-Wilson, K., Agnew, J., and Meyers, D. A. (1988). Age-related sex differences in verbal memory. *J. Clin. Psychol.* 44, 403–411. doi: 10.1002/1097-4679(198805)44:3<403::aid-jclp2270440315>3.0.co;2-0

Bolla-Wilson, K., and Bleecker, M. L. (1986). Influence of verbal intelligence, sex, age, and education on the Rey auditory verbal learning test. *Dev. Neuropsychol.* 2, 203–211. doi: 10.1080/87565648609540342

Bussè, C., Anselmi, P., Pompanin, S., Zorzi, G., Fragiacomo, F., Camporese, G., et al. (2017). Specific verbal memory measures may distinguish Alzheimer's disease from dementia with lewy bodies. *J. Alzheimers Dis.* 59, 1009–1015. doi: 10.3233/JAD-170154

Campoy, G., Castellà, J., Provencio, V., Hitch, G. J., and Baddeley, A. D. (2015). Automatic semantic encoding in verbal short-term memory: evidence from the concreteness effect. *Q. J. Exp. Psychol.* 68, 759–778. doi: 10.1080/17470218.2014.966248

Casaletto, K. B., Marx, G., Dutt, S., Neuhaus, J., Saloner, R., Kritikos, L., et al. (2017). Is "Learning" episodic memory? Distinct cognitive and neuroanatomic correlates of immediate recall during learning trials in neurologically normal aging and neurodegenerative cohorts. *Neuropsychologia* 102, 19–28. doi: 10.1016/j.neuropsychologia.2017.05.021

Chapman, R. M., Mapstone, M., Gardner, M. N., Sandoval, T. C., McCrary, J. W., Guillily, M. D., et al. (2011). Women have farther to fall: gender differences between normal elderly and Alzheimer's disease in verbal memory engender better detection of Alzheimer's disease in women. *J. Int. Neuropsychol. Soc.* 17, 654–662. doi: 10.1017/S1355617711000452

Consonni, M., Rossi, S., Cerami, C., Marcone, A., Iannaccone, S., Francesco Cappa, S., et al. (2017). Executive dysfunction affects word list recall performance: evidence from amyotrophic lateral sclerosis and other neurodegenerative diseases. *J. Neuropsychol.* 11, 74–90. doi: 10.1111/jnp.12072

Constantinidou, F., Zaganas, I., Papastefanakis, E., Kasselimis, D., Nidos, A., and Simos, P. G. (2014). Age-related decline in verbal learning is moderated by demographic factors, working memory capacity, and presence of amnestic mild

cognitive impairment. *J. Int. Neuropsychol. Soc.* 20, 822–835. doi: 10.1017/S1355617714000678

Craik, F. I. M., and McDowd, J. M. (1987). Age differences in recall and recognition. *J. Exp. Psychol. Learn. Mem. Cogn.* 13, 474–479. doi: 10.1037/0278-7393.13.3.474

Croisile, B., Ska, B., Brabant, M.-J., Duchene, A., Lepage, Y., Aimard, G., et al. (1996). Comparative study of oral and written picture description in patients with Alzheimer's disease. *Brain Lang.* 53, 1–19. doi: 10.1006/brln.1996.0033

Crossen, J. R., and Wiens, A. N. (1994). Comparison of the auditory-verbal learning test (AVLT) and california verbal learning test (CVLT) in a sample of normal subjects. *J. Clin. Exp. Neuropsychol.* 16, 190–194. doi: 10.1080/01688639408402630

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466.

de Lira, J. O., Ortiz, K. Z., Campanha, A. C., Bertolucci, P. H. F., and Minett, T. S. C. (2011). Microlinguistic aspects of the oral narrative in patients with Alzheimer's disease. *Int. Psychogeriatr.* 23, 404–412. doi: 10.1017/S1041610210001092

Delis, D. C., Kramer, J. H., Kaplan, E., and Thompkins, B. A. O. (1987). *CVLT: California Verbal Learning Test-Adult Version: Manual*. San Antonio, Texas: Psychological Corporation. doi: 10.1017/s1041610210001092

Desgranges, B., and Eustache, F. (2011). Les conceptions de la mémoire déclarative d'Endel Tulving et leurs conséquences actuelles. *Rev. Neuropsychol.* 3:94. doi: 10.3917/rne.032.0094

Desmette, D., Hupet, M., Schelstraete, M.-A., and Van der Linden, M. (1995). Adaptation en langue française du « reading span test » de daneman et carpenter (1980). *Année Psychol.* 95, 459–482. doi: 10.3406/psy.1995.28842

Duff, K., Schoenberg, M., Scott, J., and Adams, R. (2005). The relationship between executive functioning and verbal and visual learning and memory. *Arch. Clin. Neuropsychol.* 20, 111–122. doi: 10.1016/j.acn.2004.03.003

Duyck, W., Desmet, T., Verbeke, L. P., and Brysbaert, M. (2004). WordGen: a tool for word selection and nonword generation in Dutch, English, German, and French. *Behav. Res. Methods Instrum. Comput.* 36, 488–499. doi: 10.3758/bf03195595

Eckerström, C., Olsson, E., Bjerke, M., Malmgren, H., Edman, Å, Wallin, A., et al. (2013). A combination of neuropsychological, neuroimaging, and cerebrospinal fluid markers predicts conversion from mild cognitive impairment to dementia. *J. Alzheimers Dis.* 36, 421–431. doi: 10.3233/JAD-122440

Estévez-González, A., Kulisevsky, J., Boltes, A., Otermín, P., and García-Sánchez, C. (2003). Rey verbal learning test is a useful tool for differential diagnosis in the preclinical phase of Alzheimer's disease: comparison with mild cognitive impairment and normal aging: rey verbal learning test. *Int. J. Geriatr. Psychiatry* 18, 1021–1028. doi: 10.1002/gps.1010

Eustache, F., Rioux, P., Desgranges, B., Marchal, G., Petit-Taboué, M.-C., Dary, M., et al. (1995). Healthy aging, memory subsystems and regional cerebral oxygen consumption. *Neuropsychologia* 33, 867–887. doi: 10.1016/0028-3932(95)00021-T

Gale, S. D., Baxter, L., Connor, D. J., Herring, A., and Comer, J. (2007). Sex differences on the rey auditory verbal learning test and the brief visuospatial memory test-revised in the elderly: normative data in 172 participants. *J. Clin. Exp. Neuropsychol.* 29, 561–567. doi: 10.1080/13803390600864760

Gale, S. D., Baxter, L., and Thompson, J. (2016). Greater memory impairment in dementing females than males relative to sex-matched healthy controls. *J. Clin. Exp. Neuropsychol.* 38, 527–533. doi: 10.1080/13803395.2015.1132298

Gavett, B. E., and Horwitz, J. E. (2012). Immediate list recall as a measure of short-term episodic memory: insights from the serial position effect and item response theory. *Arch. Clin. Neuropsychol.* 27, 125–135. doi: 10.1093/arclin/acr104

Geffen, G., Moar, K. J., O'hanlon, A. P., Clark, C. R., and Geffen, L. B. (1990). Performance measures of 16– to 86-year-old males and females on

the auditory verbal learning test. *Clin. Neuropsychol.* 4, 45–63. doi: 10.1080/13854049008401496

Goryawala, M., Zhou, Q., Barker, W., Loewenstein, D. A., Duara, R., and Adjouadi, M. (2015). Inclusion of neuropsychological scores in atrophy models improves diagnostic classification of Alzheimer's disease and mild cognitive impairment. *Comput. Intell. Neurosci.* 2015, 1–14. doi: 10.1155/2015/865265

Griffin, J. W., John, S. E., Adams, J. W., Bussell, C. A., Saurman, J. L., and Gavett, B. E. (2017). The effects of age on the learning and forgetting of primacy, middle, and recency components of a multi-trial word list. *J. Clin. Exp. Neuropsychol.* 39, 900–912. doi: 10.1080/13803395.2017.1278746

Halekoh, U., and Højsgaard, S. (2014). A kenward-roger approximation and parametric bootstrap methods for tests in linear mixed models – The R package pbkrtest. *J. Stat. Softw.* 59, 9024. doi: 10.18637/jss.v059.i09

Herlitz, A., and Viitanen, M. (1991). Semantic organization and verbal episodic memory in patients with mild and moderate Alzheimer's disease. *J. Clin. Exp. Neuropsychol.* 13, 559–574. doi: 10.1080/01688639108401071

Herrmann, D. J. (1987). Task appropriateness of mnemonic techniques. *Percept. Mot. Skills* 64, 171–178. doi: 10.2466/pms.1987.64.1.171

Ivnik, R. J., Malec, J. F., Tangalos, E. G., Petersen, R. C., Kokmen, E., and Kurland, L. T. (1990). The auditory-verbal learning test. *Psychol. Assess.* 2, 304–312.

Kelley, K. (2019). *R package "MBESS"* (Version 4.0.6). Available at: https://cran.r-project.org/package=MBESS (accessed July 10, 2019).

Kirkby, K. C., Hennessy, M. J., Montgomery, I. M., and Daniels, B. A. (1995). Amnesia following gastroscopy with midazolam: a comparison in two age groups. *J. Psychopharmacol. (Oxf.)* 9, 32–37. doi: 10.1177/026988119500900106

Kramer, J. H., Delis, D. C., and Daniel, M. (1988). Sex differences in verbal learning. *J. Clin. Psychol.* 44, 10.

Kramer, J. H., Delis, D. C., Kaplan, E., O'Donnell, L., and Prifitera, A. (1997). Developmental sex differences in verbal learning. *Neuropsychology* 11, 577–584. doi: 10.1037/0894-4105.11.4.577

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest Package: tests in linear mixed effects models. *J. Stat. Softw.* 82, doi: 10.18637/jss.v082.i13

Lenth, R. V. (2020). *Package 'emmeans'*. Available at: https://cran.r-project.org/web/packages/emmeans/emmeans.pdf (accessed March 10, 2020).

Lezak, M. D. (1983). *Neuropsychological Assessment*, 2nd Edn. Oxford: Oxford University Press.

Li, K., Chan, W., Doody, R. S., Quinn, J., Luo, S., and Alzheimer's Disease Neuroimaging Initiative. (2017). Prediction of Conversion to Alzheimer's disease with longitudinal measures and time-to-event data. *J. Alzheimers Dis.* 58, 361–371. doi: 10.3233/JAD-161201

Litvan, I., Mohr, E., Williams, J., Gomez, C., and Chase, T. N. (1991). Differential memory and executive functions in demented patients with Parkinson's and Alzheimer's disease. *J. Neurol. Neurosurg. Psychiatry* 54, 25–29. doi: 10.1136/jnnp.54.1.25

Long, J. A. (2019). *Interactions: Comprehensive, User-Friendly Toolkit for Probing Interactions*. Available at: https://cran.r-project.org/package=interactions (accessed September 12, 2019).

Magalhães, S. S., and Hamdan, A. C. (2010). The rey auditory verbal learning test: normative data for the brazilian population and analysis of the influence of demographic variables. *Psychol. Neurosci.* 3, 85–91. doi: 10.3922/j.psns.2010.1.011

Maitland, S. B., Herlitz, A., Nyberg, L., Bäckman, L., and Nilsson, L.-G. (2004). Selective sex differences in declarative memory. *Mem. Cognit.* 32, 1160–1169. doi: 10.3758/BF03196889

Malloy-Diniz, L. F., and Parreira, V. A. (2007). The rey auditory-verbal learning test: applicability for the brazilian elderly population Teste de aprendizagem auditivo-verbal de Rey: aplicabilidade na população idosa brasileira. *Rev Bras Psiquiatr* 6, 324–329.

Marra, C., Silveri, M. C., and Gainotti, G. (2000). Predictors of cognitive decline in the early stage of probable Alzheimer's disease. *Dement. Geriatr. Cogn. Disord.* 11, 212–218. doi: 10.1159/000017239

Mazoyer, B., Mellet, E., Perchey, G., Zago, L., Crivello, F., Jobard, G., et al. (2016). BIL&GIN: a neuroimaging, cognitive, behavioral, and genetic database for the study of human brain lateralization. *NeuroImage* 124, 1225–1231. doi: 10.1016/j.neuroimage.2015.02.071

Mazoyer, B., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., et al. (1993). The cortical representation of speech. *J. Cogn. Neurosci.* 5, 467–479. doi: 10.1162/jocn.1993.5.4.467

McMinn, M. R., Wiens, A. N., and Crossen, J. R. (1988). Rey auditory-verbal learning test: development of norms for healthy young adults. *Clin. Neuropsychol.* 2, 67–87. doi: 10.1080/13854048808520087

Meijs, C. J. C., Hurks, P., Rozendaal, N., and Jolles, J. (2013). Serial and subjective clustering on a verbal learning test (VLT) in children aged 5–15: the nature of subjective clustering. *Child Neuropsychol.* 19, 385–399. doi: 10.1080/09297049.2012.670215

Mellet, E., Jobard, G., Zago, L., Crivello, F., Petit, L., Joliot, M., et al. (2013). Relationships between hand laterality and verbal and spatial skills in 436 healthy adults balanced for handedness. *Laterality Asymmetries Body Brain Cogn.* 19, 383–404. doi: 10.1080/1357650X.2013.796965

Messinis, L., Tsakona, I., Malefaki, S., and Papathanasopoulos, P. (2007). Normative data and discriminant validity of Rey's verbal learning test for the greek adult population. *Arch. Clin. Neuropsychol.* 22, 739–752. doi: 10.1016/j.acn.2007.06.002

Mitrushina, M., Satz, P., Chervinsky, A., and D'Elia, L. (1991). Performance of four age groups of normal elderly on the rey auditory-verbal learning test. *J. Clin. Psychol.* 47, 351–357. doi: 10.1002/1097-4679(199105)47:3<351::aid-jclp2270470305>3.0.co;2-s

Moffot, A. P. R., O'Carroll, R. E., Bennie, J., Carroll, S., Dick, H., Ebmeier, K. P., et al. (1994). Diurnal variation of mood and neuropsychological function in major depression with melancholia. *J. Affect. Disord.* 32, 257–269. doi: 10.1016/0165-0327(94)90090-6

Moradi, E., Hallikainen, I., Hänninen, T., and Tohka, J. (2017). Rey's auditory verbal learning test scores can be predicted from whole brain MRI in Alzheimer's disease. *NeuroImage Clin.* 13, 415–427. doi: 10.1016/j.nicl.2016.12.011

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., et al. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* 39, 1159–1165.

Morrison, R. L., Pei, H., Novak, G., Kaufer, D. I., Welsh-Bohmer, K. A., Ruhmel, S., et al. (2018). A computerized, self-administered test of verbal episodic memory in elderly patients with mild cognitive impairment and healthy participants: a randomized, crossover, validation study. *Alzheimers Dement. Diagn. Assess. Dis. Monit.* 10, 647–656. doi: 10.1016/j.dadm.2018.08.010

Mortensen, E. L., and Gade, A. (1993). On the relation between demographic variables and neuropsychological test performance. *Scand. J. Psychol.* 34, 305–317. doi: 10.1111/j.1467-9450.1993.tb01127.x

Nakagawa, S., Johnson, P. C. D., and Schielzeth, H. (2017). The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *J. R. Soc. Interface* 14:20170213.

Nebel, R. A., Aggarwal, N. T., Barnes, L. L., Gallagher, A., Goldstein, J. M., Kantarci, K., et al. (2018). Understanding the impact of sex and gender in Alzheimer's disease: a call to action. *Alzheimers Dement.* 14, 1171–1183. doi: 10.1016/j.jalz.2018.04.008

Park, D. C., and Festini, S. B. (2017). Theories of memory and aging: a look at the past and a glimpse of the future. *J. Gerontol. B. Psychol. Sci. Soc. Sci.* 72, 82–90. doi: 10.1093/geronb/gbw066

Park, L. Q., Gross, A. L., McLaren, D. G., Pa, J., Johnson, J. K., Mitchell, M., et al. (2012). Confirmatory factor analysis of the ADNI neuropsychological battery. *Brain Imaging Behav.* 6, 528–539. doi: 10.1007/s11682-012-9190-3

Poreh, A. (2005). Analysis of mean learning of normal participants on the rey auditory-verbal learning test. *Psychol. Assess.* 17, 191–199. doi: 10.1037/1040-3590.17.2.191

Preacher, K. J., and Kelley, K. (2011). Effect size measures for mediation models: quantitative strategies for communicating indirect effects. *Psychol. Methods* 16, 93–115. doi: 10.1037/a0022658

Putcha, D., Brickhouse, M., Wolk, D. A., and Dickerson, B. C. (2019). Fractionating the rey auditory verbal learning test: distinct roles of large-scale cortical networks in prodromal Alzheimer's disease. *Neuropsychologia* 129, 83–92. doi: 10.1016/j.neuropsychologia.2019.03.015

Rey, A. (1970). *L'examen Clinique En Psychologie. [The clinical examination in psychology.]*, 3e édn. Paris: Presses Universitaries De France.

Ricci, M., Graef, S., Blundo, C., and Miller, L. A. (2012). Using the rey auditory verbal learning test (RAVLT) to differentiate Alzheimer's dementia and behavioural variant fronto-temporal dementia. *Clin. Neuropsychol.* 26, 926–941. doi: 10.1080/13854046.2012.704073

Roediger, H. L. (1990). Implicit memory: retention without remembering. *Am. Psychol.* 45:1043. doi: 10.1037/0003-066x.45.9.1043

Saury, J.-M., and Emanuelson, I. (2017). Neuropsychological assessment of hippocampal integrity. *Appl. Neuropsychol. Adult* 24, 140–151. doi: 10.1080/23279095.2015.1113536

Sederberg, P. B., Miller, J. F., Howard, M. W., and Kahana, M. J. (2010). The temporal contiguity effect predicts episodic memory performance. *Mem. Cognit.* 38, 689–699. doi: 10.3758/MC.38.6.689

Selnes, O. A., Jacobson, L., Machado, A. M., Becker, J. T., Wesch, J., Miller, E. N., et al. (1991). Normative data for a brief neuropsychological screening battery. *Percept. Mot. Skills* 73, 539–550. doi: 10.2466/pms.1991.73.2.539

Shaywitz, B. A., Shaywltz, S. E., Pugh, K. R., Constable, R. T., Skudlarski, P., Fulbright, R. K., et al. (1995). Sex differences in the functional organization of the brain for language. *Nature* 373:607.

Spreen, O., Spreen, P. O. P. O., Strauss, E., and Strauss, P. O. P. E. (1998). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary.* Oxford, MI: Oxford University Press.

Stricker, J. L., Brown, G. G., Wixted, J., Baldo, J. V., and Delis, D. C. (2002). New semantic and serial clustering indices for the California Verbal Learning Test–Second Edition: background, rationale, and formulae. *J. Int. Neuropsychol. Soc.* 8, 425–435. doi: 10.1017/S1355617702813224

Sudo, F. K., de Souza, A. S., Drummond, C., Assuncao, N., Teldeschi, A., Oliveira, N., et al. (2019). Inter-method and anatomical correlates of episodic memory tests in the Alzheimer's Disease spectrum. *PLoS ONE* 14:e0223731. doi: 10.1371/journal.pone.0223731

Sunderaraman, P., Blumen, H. M., DeMatteo, D., Apa, Z. L., and Cosentino, S. (2013). Task demand influences relationships among sex, clustering strategy, and recall: 16-word versus 9-word list learning tests. *Cogn. Behav. Neurol.* 26, 78–84. doi: 10.1097/WNN.0b013e31829de450

Sundermann, E. E., Biegon, A., Rubin, L. H., Lipton, R. B., Landau, S., Maki, P. M., et al. (2017). Does the female advantage in verbal memory contribute to underestimating Alzheimer's disease pathology in women versus men? *J. Alzheimers Dis.* 56, 947–957. doi: 10.3233/JAD-160716

Teruya, L. C., Ortiz, K. Z., and Minett, T. S. C. (2009). Performance of normal adults on Rey Auditory Learning Test: a pilot study. *Arq. Neuropsiquiatr.* 67, 224–228. doi: 10.1590/S0004-282X2009000200010

Tulving, E. (1985). Ebbinghaus's memory: what did he learn and remember? *J. Exp. Psychol. Learn. Mem. Cogn.* 11, 485–490. doi: 10.1037/0278-7393.11.3.485

Tulving, E. (1995). "Organization of memory: quo vadis," in *The Cognitive Neurosciences*, ed. M. S. Gazzaniga (Cambridge: The MIT Press).

Tulving, E. (2002). Episodic memory: from mind to brain. *Annu. Rev. Psychol.* 53, 1–25. doi: 10.1146/annurev.psych.53.100901.135114

Tuokko, H., Griffith, L. E., Simard, M., and Taler, V. (2017). Cognitive measures in the Canadian Longitudinal Study on Aging. *Clin. Neuropsychol.* 31, 233–250. doi: 10.1080/13854046.2016.1254279

Unsworth, N., Miller, A. L., and Robison, M. K. (2019). Individual differences in encoding strategies and free recall dynamics. *Q. J. Exp. Psychol.* doi: 10.1177/1747021819847441

Uttl, B. (2005). Measurement of individual differences. *Psychol. Sci.* 16, 460–467.

Vakil, E., and Blachstein, H. (1994). A supplementary measure in the Rey AVLT for assessing incidental learning of temporal order. *J. Clin. Psychol.* 50, 240–245. doi: 10.1002/1097-4679(199403)50:2<240::aid-jclp2270500215>3.0.co;2-5

Van Der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., and Jolles, J. (2005). Rey's verbal learning test: normative data for 1855 healthy participants aged 24–81 years and the influence of age, sex, education, and mode of presentation. *J. Int. Neuropsychol. Soc.* 11, 290–302. doi: 10.1017/S1355617705050344

Van der Linden, M. (2004). "Fondements théoriques et méthodologiques," in *L'évaluation des Troubles de la Mémoire Épisodique (Marseille)* (Marseille: Solal).

Venables, W. N., and Ripley, B. D. (2002). *Modern Applied Statistics with S*, 4th Edn. New York, NY: Springer.

Verma, M., and Howard, R. J. (2012). Semantic memory and language dysfunction in early Alzheimer's disease: a review. *Int. J. Geriatr. Psychiatry* 27, 1209–1217. doi: 10.1002/gps.3766

Wen, Z., and Fan, X. (2015). Monotonicity of effect sizes: questioning kappa-squared as mediation effect size measure. *Psychol. Methods* 20, 193–203. doi: 10.1037/met0000029

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* Berlin: Springer.

Worthen, J. B., and Hunt, R. R. (2017). "Mnemonic techniques: underlying processes and practical applications," in *Learning and Memory: A Comprehensive Reference* (Amsterdam: Elsevier), 515–527. doi: 10.1016/b978-0-12-809324-5.21063-8

Ziegler, J. C., Besson, M., Jacobs, A. M., Nazir, T. A., and Carr, T. H. (1997). Word, pseudoword, and nonword processing: a multitask comparison using event-related brain potentials. *J. Cogn. Neurosci.* 9, 758–775. doi: 10.1162/jocn.1997.9.6.758