Published in final edited form as:
Cogn Behav Neurol. 2013 June ; 26(2): 78–84. doi:10.1097/WNN.0b013e31829de450.

Task Demand Influences Relationships Among Sex, Clustering Strategy, and Recall: 16-Word Versus 9-Word List Learning Tests

Preeti Sunderaraman, BA[*], **Helena M. Blumen, PhD**[†], **David DeMatteo, PhD**[*], **Zoltan Apa, BA**[‡], and **Stephanie Cosentino, PhD**[§]

[*]Department of Psychology, Drexel University, Philadelphia, PA

[†]Albert Einstein College of Medicine, Bronx, NY

[‡]School of Nursing, College of Physicians and Surgeons, Columbia University, New York, NY

[§]The Taub Institute for Research on Alzheimer's disease and the Aging Brain, The Gertrude H. Sergievsky Center, and the Department of Neurology, Columbia University Medical Center, New York, NY

Abstract

Objective—We compared the relationships among sex, clustering strategy, and recall across different task demands using the 16-word California Verbal Learning Test–Second Edition (CVLT-II) and the 9-word Philadelphia (repeatable) Verbal Learning Test (PrVLT).

Background—Women generally score higher than men on verbal memory tasks, possibly because women tend to use semantic clustering. This sex difference has been established via word-list learning tests such as the CVLT-II.

Methods—In a retrospective between-group study, we compared how 2 separate groups of cognitively healthy older adults performed on a longer and a shorter verbal learning test. The group completing the CVLT-II had 36 women and 26 men; the group completing the PrVLT had 27 women and 21 men.

Results—Overall, multiple regression analyses revealed that semantic clustering was significantly associated with total recall on both tests' lists ($P < 0.001$). Sex differences in recall and semantic clustering diminished with the shorter PrVLT word list.

Conclusions—Semantic clustering uniquely influenced recall on both the longer and shorter word lists. However, serial clustering and sex influenced recall depending on the length of the

Corresponding Author: Stephanie Cosentino, PhD, Assistant Professor of Neuropsychology, Columbia University Medical Center, 630 West 168th Street, P&S Box 16, New York, NY 10032, Phone: (212) 342 0289, Fax: (212) 342-1838, sc2460@columbia.edu. Please send the page proof to both Dr. Cosentino and the first author: Preeti Sunderaraman, preetis.ps@gmail.com, 917 584-8024

P.S., S.C., and H.M.B. conceived of and designed the study. S.C., H.M.B., P.S., and Z.A. collected the data. P.S., S.C., H M.B., and D.DeM. analyzed and interpreted the data.

The authors declare no conflicts of interest.

[Statement of reader benefit:] Organizing information based on meaning enhances learning regardless of task difficulty; women's tendency to use this strategy contributes to their performance advantage at both high and low levels of information load.

word list (ie, the task demand). These findings suggest a complex nonlinear relationship among verbal memory, clustering strategies, and task demand.

Keywords

verbal memory; semantic clustering; sex differences; task demand; older adults

The brain's organizational methods during encoding and recall can influence how successfully we are able to recall words. Semantic clustering and serial clustering are the 2 main methods for organizing word recall on categorized word-list learning tests such as the California Verbal Learning Test, Second Edition (CVLT-II) (Delis et al, 2000; Delis and Fridlund, 2000) and the Philadelphia (repeatable) Verbal Learning Test (PrVLT) (Price et al, 2009). Semantic clustering is the tendency to recall words according to their category membership; serial clustering is the tendency to recall words in the same order as they were originally presented. Semantic clustering has been associated with better recall performance than serial clustering (Delis et al, 2000; Donders, 2008; Paolo et al, 1997).

Studies evaluating sex differences in memory tasks have found that women tend to perform better than men on most recall tests (Herlitz, Nilsson, and Backman, 1997). In general, sex differences are attributed to semantic clustering when the to-be-remembered information can be organized semantically, as is the case on most word-list learning tests, including the CVLT and PrVLT. In fact, the greater success of semantic than serial clustering may account for sex differences in verbal recall performance on semantically organizable lists. Many studies have shown that women recall more words than men on most word-list learning tests, including the CVLT-II, and that this is likely a function of their greater tendency to cluster words semantically during recall (Andreano and Cahill, 2009).

Most studies that have examined the relationships among sex, clustering method, and recall performance have used 12- to 16-word-list tests such as the CVLT-II, the Hopkins Verbal Learning Test (Brandt and Benedict, 2001), the Buschke Selective Reminding Test (Buschke, 1973; Buschke and Fuld, 1974), and the Rey Auditory Verbal Learning Test (Schmidt, 1996). The relationship between clustering method and recall performance on a shorter word-list learning test such as the 9-word PrVLT (Price et al, 2009) had not been investigated until now, leaving open the question of whether semantic strategies were advantageous for recall at any level of task demand, or whether their utility lay in organizing stimuli under higher task demand. Also unknown was the extent to which sex differences in clustering and recall would persist at lower task demand.

In the current study, we presumed that a 9-word list would have a lower task demand than a 16-word list. One possibility was that serial clustering would be as effective as semantic clustering for recalling a shorter list if the use of a semantic strategy was based on the need to organize information that exceeded a certain memory demand. Another possibility was that we would confirm Craik and Lockhart's (1972) postulate that because of the deeper nature of encoding, semantic clustering might optimize recall regardless of task demand.

We studied the relationships among sex, clustering strategy, and recall on verbal memory tests under the high task demand of the CVLT-II and the lower task demand of the PrVLT. The well-established CVLT-II is used to test learning and recall in people across the lifespan, with and without neurologic deficits (Lezak, Howieson, and Loring, 2004). The newer PrVLT is used to characterize memory deficits in people with mild cognitive impairment, Alzheimer disease, vascular dementia, and leukoaraiosis (Price et al, 2009; Libon et al, 1998; Libon et al, 2010).

The main difference between the 2 tests is the number of words to be learned: 16 for the CVLT-II vs 9 for the PrVLT. Otherwise, the 2 tests share much in common. They have similar instructions and administration. Both involve 5 learning trials, 1 interference trial, a short-delay free and cued recall, a long-delay free and cued recall, and a yes-no recognition task. The 2 tests also have some similarities in their semantic categories and word lists. The CVLT-II has 4 semantic categories: vegetables, ways of traveling, furniture, and animals; the PrVLT has 3 categories: fruits, furniture, and tools. The design of the 2 tasks' word lists is broadly comparable, with both including natural and artificial categories, and both excluding prototypical words like "apple," "bed," and "hammer." Thus, while the 2 lists differ in their specific content, the general content is similar, leaving the number of words as the most prominent difference.

The goal of this study was to examine whether the expected effects of semantic clustering and sex on a long word-list learning test would hold for a shorter word-list learning test. Analogous to the long word lists, we wanted to determine whether women's recall advantage over men could be entirely accounted for by women's greater use of semantic clustering. In line with existing studies, we hypothesized that on a 16-word list, semantic clustering would be more beneficial to recall than would serial clustering, and that women would use more semantic clustering and demonstrate higher total immediate recall than men. In contrast, on the shorter 9-word list, we adopted the null hypothesis that there would be no sex differences in clustering strategies and total immediate recall, and that we would not find an association among sex, clustering strategy, and recall.

METHODS

Participants

We performed a retrospective analysis on data from 2 previous studies performed by our group at the Alzheimer's Disease Research Center at Columbia University Medical Center. We analyzed the findings of reports on 2 separate populations of cognitively healthy older adults: 36 women and 26 men who had completed the CVLT-II (Stern et al, 2011), and 27 women and 21 men who had completed the PrVLT (Cosentino et al, 2011).

Both studies had been performed at around the same time. As described below, both had recruited and screened participants at Columbia using similar exclusion and inclusion criteria and comparable screening measures.

The participants in the CVLT-II study had been recruited via fliers, internet ads, and visits to local senior centers. The participants in the PrVLT study had been recruited from the database of healthy controls available through the Alzheimer's Disease Research Center, as well as at local senior centers and through marketing mailings targeting a diverse population of New York City elders.

For both studies, candidates had first been screened for their demographic, general medical, neurologic, and psychiatric history and for dementia through a phone questionnaire and, if available, records from participation in earlier studies. Candidates were excluded if they had been diagnosed with any major neurologic or psychiatric condition or if they had suffered a head injury that caused them to lose consciousness. Once recruited, participants had been disqualified if the examiner found them to have significant sensory or motor abnormalities that might interfere with cognitive testing, eg, vision impairment that was not corrected by eyeglasses, hearing loss that was not corrected by a hearing aid, or difficulty holding and using a writing utensil.

For both studies, once participants were recruited they underwent a full neuropsychological test battery. We will focus here on the tests that ruled out dementia and assessed intellectual function and verbal fluency. Because the 2 studies had been performed independently, they used somewhat different measures to collect these data.

As a test for dementia, the participants in the CVLT-II study completed the Mattis Dementia Rating Scale (Jurica et al, 2004). To stay in the study, they had to have a total score of 135 or higher; anything lower indicated cognitive impairment. The participants in the PrVLT study were tested for dementia with the Mini-Mental State Examination (Folstein et al, 1975). They stayed in the study if they scored at least 24/30; anything lower indicated significant cognitive impairment.

As a measure of intellectual function, the participants in both groups were given the Wechsler Test of Adult Reading (Wechsler, 2001). As a measure of verbal fluency, all participants were given the Controlled Oral Word Association Test (Benton and Hamsher, 1976); however, the CVLT-II group was given the "CFL" form, and the PrVLT group was given the "FAS" form.

The Stern et al (2011) and Cosentino et al (2011) studies had been approved by the Columbia University Institutional Review Board, and the participants had given informed consent and been paid for taking part. No further approval was required for our study.

Procedures

Each participant completed only the CVLT-II or only the PrVLT. As mentioned, the major difference between the 2 tasks is their length (task demand): The 16-word CVLT-II has 4 semantic categories, each category consisting of 4 words; the 9-word PrVLT has 3 semantic categories, each consisting of 3 words.

Although each participant had completed the entire task, we chose to focus our analysis only on the total scores for the 5 immediate recall trials. We did not study the data for the delayed recall trials. Delayed recall trials can be confounded by participants' exposure to the interference list, whose semantic categories are similar to the target list and thus can influence participants' organizational strategy. Clustering strategy can also be influenced by exposure to semantic cuing after the short-delay trial (Woo and Schmitter-Edgecombe, 2009).

We evaluated 2 scores: total immediate recall across trials 1 to 5, and total clustering strategy scores across trials. Clustering strategy scores comprised semantic and *serial bidirectional scores*. Unlike scores for *serial forward* (recalling words in the same consecutive forward direction in which they were presented) or *serial backward* (recalling words in the same consecutive backward direction in which they were presented) considered in isolation, serial bidirectional scores combine serial forward and serial backward scores to capture any serial approach to recalling the list.

We adjusted both the semantic and serial clustering scores for the number of clusters that would have been recalled at chance level for a given number of words recalled (Stricker et al, 2002). For each of the 5 immediate recall trials in the PrVLT, we calculated the semantic and serial clustering chance-adjusted scores, based on the formula presented in the CVLT-II manual (Delis et al, 2000). For the CVLT-II and the PrVLT, the possible semantic chance-adjusted scores ranged from −3 to 9 and from −2 to 4, respectively, and the possible serial chance-adjusted scores ranged from −1.88 to +13.13 and from −1.78 to 6.22, respectively. Given the differences in list length, the total possible number of semantic and serial clustering scores differed across test.

Statistical Analysis

We used PASW Statistics version 18.0 for all analyses. To compare the differences between women and men by demographic characteristics and neuropsychological variables, we conducted independent samples *t* tests separately for each sample. We conducted 3 independent samples *t* tests to examine sex differences in total immediate recall, semantic clustering, and serial clustering.

We also conducted 2 simultaneous linear regression analyses for each word-list test to assess the relationships among clustering strategy, sex, and total immediate recall. Our first regression analysis evaluated the relationship between semantic clustering and serial clustering (as the predictor variables) and total immediate recall (as the outcome variable). Our second regression analysis evaluated the relationship between semantic clustering, serial clustering, and sex (as the predictor variables) and total immediate recall (as the outcome variable). Formal power analysis (with $\alpha = 0.05$, effect size = medium, and power = 0.80) revealed that a regression-based analysis with 2 and 3 predictor variables required sample sizes of 67 and 76 participants, respectively (Cohen, 1992).

RESULTS

Demographic Characteristics

Our CVLT-II and PrVLT study groups were similar in age (mean ages were 66.07 and 67.85 years for the CVLT-II and PrVLT groups, respectively) and education (means were 15.97 and 15.54 years for the CVLT-II and the PrVLT groups, respectively) (Table 1). Although the 2 study groups had differences in ethnicity, chi-square analysis did not reveal any significant effects. The 2 study groups were also comparable in their estimated intelligence quotient based on the Wechsler Test of Adult Reading (Wechsler, 2001). Within each study group, women and men performed similarly on the Controlled Oral Word Association Test, as assessed with the “CFL” form (Benton et al, 1994) in the CVLT-II group and the “FAS” form (Gladsjo et al, 1999) in the PrVLT group. Finally, cognitive screening measures were comparable in women and men, as assessed with the Mattis Dementia Rating Scale (Mattis, 1976) in the CVLT-II group and with the Mini-Mental State Examination (Folstein et al, 1975) in the PrVLT group.

CVLT-II Results

For the CVLT-II group, an independent samples *t* test showed that total immediate recall was greater in women (mean [M] = 50.92, standard deviation [SD] = 8.81) than men (M = 40.81, SD = 7.50), $t(60) = 4.74$, $P < 0.001$. Semantic clustering was also greater in women (M = 1.08, *SD* = 1.86) than men (M = −0.02, SD = 0.53), $t(60) = 2.92$, $P < 0.01$, while serial clustering was similar in women and men, $t(60) = -0.35$, $P = 0.73$. Table 2 shows the means and standard deviations of total immediate recall and the clustering strategies for each sex.

Correlations between the 2 predictor variables semantic clustering and serial clustering ($r = -0.56$) did not indicate high multicollinearity (Table 3). We found a significant positive correlation between semantic clustering and total immediate recall ($r = 0.52$, $P < 0.001$). However, we also found a positive but nonsignificant correlation between serial clustering and total immediate recall ($r = 0.06$, $P = 0.32$).

The first regression analysis (model 1) revealed that when both clustering strategies were used as predictors, approximately 43% of the variance in total immediate recall could be predicted. About 45% of variance in total immediate recall was uniquely predicted from semantic clustering (while serial clustering was statistically controlled), and about 18% of variance in total immediate recall was uniquely predicted from serial clustering. As for the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

relation between semantic clustering and total immediate recall, because the semipartial correlations were higher than the zero-order correlations, we concluded that serial clustering might be acting as a suppressor variable (ie, a third variable, which, if included in the regression equation, would increase the relation between the first and second variables). Thus, the unique predictive contribution of semantic clustering to total immediate recall was increased when serial clustering was included in the regression analyses.

The results of the overall multiple regression using both of the clustering strategies to predict total immediate recall were significant.

The second regression analysis showed that when both semantic and serial clustering and sex were used as predictors, about 50% of the variance in total immediate recall could be predicted. About 25% of variance was uniquely predicted from semantic clustering, about 13% was uniquely predicted from serial clustering, and about 8% was uniquely predicted from sex.

The overall multiple regression using clustering strategies and sex to predict total immediate recall was significant. Specifically, the relation between sex and total immediate recall also indicated that men had a lower total score than women.

PrVLT Results

Echoing the pattern of findings for the CVLT-II, an independent samples t test for the PrVLT showed that women (M = 39.00, SD = 3.39) performed significantly better than men (M = 36.29, SD = 4.45) on total immediate recall, $t(47) = 2.43$, $P < 0.05$ (Table 2). Semantic clustering was also greater in women (M = 1.27, SD = 1.30) than men (M = 0.46, SD = 1.49), $t(47) = 2.03$, $P < 0.05$, while serial clustering was similar in women and men, $t(47) = -0.79$, $P = 0.43$.

Correlations between the 2 predictor variables semantic clustering and serial clustering ($r = -0.73$) did not indicate extremely high multicollinearity (Table 3). The first regression analysis (model 1) revealed that when both clustering strategies were used as predictors, about 31% of the variance in total immediate recall could be predicted (overall), and about 20% of variance was uniquely predicted from semantic clustering. The unique contribution of serial clustering was not significant in this model. For the overall multiple regression to predict total immediate recall from clustering strategies, $R = 0.58$, $R^2 = 0.34$, and adjusted $R^2 = 0.31$.

The overall regression was statistically significant: $F(48,2) = 11.90$, $P < 0.001$. Results of the first analysis examining the relationship between semantic clustering and total immediate recall, and between serial clustering and total immediate recall (in the absence of the predictor sex) indicated that only semantic clustering ($b = 1.88$, $\beta = 0.66$, standard error [SE] = 0.50, $t = 3.74$, $P < 0.001$) significantly predicted total immediate recall. The square of the semipartial correlation (sr^2) that estimated how much variance in total immediate recall was uniquely predicted from semantic clustering was 0.20.

The second regression analysis showed that when both of the clustering strategies and sex were used as predictors, about 33% of the variance in total immediate recall could be predicted. About 14% of the variance was uniquely predicted from semantic clustering. The unique predictions of serial clustering and sex were not significant. For the overall multiple regression to predict total immediate recall from clustering strategies and sex, $R = 0.61$, $R^2 = 0.37$, and adjusted $R^2 = 0.33$.

The overall regression was statistically significant, $F(45, 3) = 8.78$, $P < 0.001$. Compared to the relationship of serial clustering ($b = 0.31$, $\beta = 0.07$, SE = 0.75, $t = 0.41$, $P = 0.69$) and total immediate recall, and the relationship of sex ($b = -1.44$, $\beta = -0.18$, SE = 1.02, $t = -1.42$, $P = 0.16$) and total immediate recall, only the relation between semantic clustering ($b = 1.66$, $\beta = 0.58$, SE = 0.52, $t = 3.19$, $P < 0.01$) and total immediate recall was significant. Specifically, the results of the relation between sex and total immediate recall also indicated that there was no significant difference in total immediate recall scores between women and men, after clustering strategies were taken into account. The square of the semipartial correlation (sr^2) that estimated how much variance in total immediate recall was uniquely predicted from semantic clustering was 0.14.

DISCUSSION

Our major aim in this study was to understand whether varying task demand by word-list length (CVLT-II vs PrVLT) would differentially influence the relationship between clustering strategies and recall. A secondary aim was to learn whether the association among clustering strategy, recall, and sex varied as a function of task demand across the lists.

In our CVLT-II group, both semantic clustering and serial clustering strategies were independently associated with recall. We also found that serial clustering might be acting as a suppressor variable. Thus, semantic clustering was associated with total immediate recall to a greater extent when serial clustering was included in the model. In other words, there appears to be a trade-off between the clustering strategies such that the more one uses a semantic strategy, the less one uses a serial strategy, and the better the total immediate recall. Therefore, controlling for serial performance yields a stronger relationship between semantic clustering and the outcome.

Much as Kramer et al (1997) reported, we found that women had both a greater tendency to use semantic clustering and higher recall scores than men. The second regression analysis showed that all 3 predictor variables—semantic clustering, serial clustering, and sex—exerted independent effects on total immediate recall. That is, the use of specific clustering strategies did not fully account for women's higher recall scores.

In the PrVLT group, only semantic clustering was associated with total immediate recall. Neither sex nor serial clustering exerted an independent effect. Analyses of sex differences showed analogous trends to those in the CVLT-II group.

We therefore conclude that while semantic clustering, serial clustering, and sex contribute to total immediate recall on a longer word-learning list, only semantic clustering contributes to total immediate recall on a shorter list. Importantly, irrespective of the length of the word-learning list, in all of our regression models the unique predictive contribution of semantic clustering to total immediate recall remained the highest.

Previous evidence for the association between semantic clustering and improved total immediate recall comes primarily from clinical studies investigating verbal memory in patients with different psychiatric and neurologic impairments (Lezak et al, 2004) and studies finding sex differences in verbal memory (Kramer et al, 1997; Kramer et al, 2003). Clinical studies using the CVLT-II showed, for instance, that compared to matched controls, patients with left (as opposed to right) temporal lobe epilepsy and those with schizophrenia tended to perform poorly on recall performance and used less semantic clustering (Brebion et al, 2004; Hermann et al, 1987). Indirect evidence for semantic clustering having a beneficial effect on recall also comes from studies comparing total immediate recall performance in women and men. The women had higher total immediate recall scores than the men on the CVLT, possibly because the women used semantic clustering.

Our study supports Kramer and colleagues' (1997) study of the original CVLT word list showing that women use more semantic clustering than men on the CVLT and have better total immediate recall. Since the formula for the clustering indices was revised in 2002 (Delis et al 2010; Stricker et al, 2002), our study has been the first to confirm Kramer's findings. However, we also found that women achieve higher recall scores than men on the CVLT-II, independent of strategy. Notably, the effects of sex diminish at lower task demand (eg, the PrVLT), while the semantic clustering advantage remains.

Our finding that only semantic clustering was associated with total immediate recall on the 9-word list suggests that the relationship shared by sex, clustering strategy, and total immediate recall is nonlinear and is affected by task demand. It appears that semantic clustering plays a crucial role in verbal recall, regardless of task demand, and that serial clustering and sex become important as the word list lengthens. The consistent benefit of semantic clustering for recall despite task demand suggests that deeper encoding processes are useful for recall across a range of task demands and are not just implemented when task demand exceeds a certain working memory threshold. Rather, it appears that with increasing list length, the nature of the task might change such that semantic clustering could become challenging to use and participants might automatically switch to serial clustering. Viewed from this perspective, it is possible that increasing list length causes information overload, reaching the limits of working memory's capacity to hold and/or categorize information semantically, and the participant might adopt another clustering strategy by default.

The role of sex under greater task demand may reflect differences in cerebral asymmetry and differential patterns of bilateral activity. For example, studies comparing lateralized behavior in women and men have found that men tend to show greater left hemisphere dominance on verbal tasks (Guillem and Mograss, 2005; McGlone, 1980). Similarly, neuroimaging studies have found that women have greater bilateral activation than men during language processing tasks (Phillips et al, 2001). It is possible that women recall more words than men because of greater bilateral activation. However, sex differences in cerebral asymmetry and neuroimaging have been debated for several decades (eg, as cited in Wallentin, 2009), and systematic research with differing word-list lengths will be needed before such conclusions can be drawn with confidence.

With a shorter word-list, both sexes might use similar neurocognitive processes and thus might perform similarly. Another possibility with a short word list is that both sexes might activate differential neurocognitive processes (as they do on a longer word list), but both processes might be equally effective in enabling memory performance given the task demand. Put another way, women's and men's performance may be similar because of the limited ability of a short word list to detect their cognitive differences, and/or the differences might manifest themselves only beyond a certain threshold. Our study had several potential limitations. One was the use of 2 separate groups of participants, 1 completing the CVLT-II and the other completing the PrVLT. However, the 2 groups were comparable across all demographic and cognitive measures, and a within-group study would have been limited by the possibility that exposure to 1 list might have influenced participants' organizational strategy and/or given rise to practice effects on the second list.

A second potential limitation was our lack of data on younger adults; however, our participants' older age enabled us to learn about memory processes in the context of cognitively normal aging. Researchers planning future studies of the relationships among sex, task demand, clustering, and recall should consider recruiting a larger, more ethnically diverse sample of older and younger adults and would ideally include biological measures such as estrogen, luteinizing hormone, and triglycerides, all of which are known to influence recall performance (Hyde et al, 2010; Kramer et al, 2003; Sims et al, 2008).

A final limitation of our study might have been low power. As noted earlier, our planned regression analysis with 2 and 3 predictor variables would have required sample sizes of 67 and 76 participants, respectively (Cohen, 1992). Our models for the relationship between clustering strategies and recall for the CVLT-II were relatively well powered with 62 participants; however, our models for the role of sex might have been underpowered, especially for the shorter PrVLT word list, which only 48 participants completed. However, our effect sizes (sr^2 unique) for the PrVLT were correspondingly lower than those for the CVLT-II (see Table 3), suggesting to some extent that the PrVLT findings reflected the nature of the task rather than power considerations per se. Future studies using larger samples might be needed to replicate this finding.

In summary, our study adds to understanding of the relationships among sex, clustering strategy, and total immediate recall. Importantly, the study highlights findings that semantic clustering appears better than serial clustering for recall across different task demands, and that women tend to use semantic clustering to a greater extent than men, especially for longer word-learning lists. Our results also suggest that, independent of clustering strategy, sex has an important effect on verbal recall at high task demand, an effect that does not arise at lower task demand.

Future studies of older adults may evaluate similar associations between semantic clustering and total immediate recall of very short (eg, 5-word) or very long (eg, 25-word) list lengths. It would also be worthwhile to explore whether women and men have a set point in list length at which they start to use semantic clustering. The outcome of such a study might help us learn whether, at longer list lengths, women and men continue to display the same pattern of performance that we saw on the CVLT-II, or whether their differences in clustering and total immediate recall become more exaggerated or, instead, disappear. Future research can also test whether specific clinical populations have similar associations among sex, clustering strategy, and recall. Such studies might help explain the neurologic basis not only of the advantages conferred by semantic clustering, but also of sex differences in verbal memory and clustering at higher task demand.

Acknowledgments

The authors thank Adam M. Brickman, PhD, for his helpful comments on the manuscript.

Glossary

CVLT-II	California Verbal Learning Test, Second Edition
M	mean
PrVLT	Philadelphia (repeatable) Verbal Learning Test
SD	standard deviation
SE	standard error

REFERENCES

Andreano JM, Cahill L. Sex influences on the neurobiology of learning and memory. Learn Mem. 2009; 16:248–266. [PubMed: 19318467]

Benton, AL.; Hamsher, KD. Multilingual Aphasia Examination. Iowa City, IA: University of Iowa; 1976.

Benton, AL.; Hamsher, KD.; Sivan, AB. Multilingual Aphasia Examination. 3rd ed. San Antonio, TX: Psychological Corporation; 1994.

Brandt, J.; Benedict, RHB. Hopkins Verbal Learning Test--Revised:Professional Manual. Lutz, FL: Psychological Assessment Resources; 2001.

Brebion G, David AS, Jones H, et al. Semantic organization and verbal memory efficiency in patients with schizophrenia. Neuropsychology. 2004; 18:378–383. [PubMed: 15099160]

Buschke H. Selective reminding for analysis of memory and learning. J Verbal Learning Verbal Behav. 1973; 12:543–550.

Buschke H, Fuld PA. Evaluating storage, retention and retrieval in disordered memory and learning. Neurology. 1974; 24:1019–1025. [PubMed: 4473151]

Cohen J. A power primer. Psychol Bull. 1992; 112:155–159. [PubMed: 19565683]

Cosentino S, Metcalfe J, Holmes B, et al. Finding the self in metacognitive evaluations: metamemory and agency in nondemented elders. Neuropsychology. 2011; 25:602–612. [PubMed: 21728426]

Craik FIM, Lockhart RS. Levels of processing: a framework for memory research. J Verbal Learning Verbal Behav. 1972; 11:671–684.

Delis DC, Fine EM, Stricker JL, et al. Comparison of the traditional recall-based versus a new list-based method for computing semantic clustering on the California Verbal Learning Test: evidence from Alzheimer's disease. Clin Neuropsychol. 2010; 24:70–79. [PubMed: 19894184]

Delis, DC.; Fridlund, AJ. CVLT-II Comprehensive Scoring Program. San Antonio, TX: The Psychological Corporation; 2000.

Delis, DC.; Kramer, JH.; Kaplan, E., et al. California Verbal Learning Test (CVLT) — Second Edition. San Antonio, TX: The Psychological Corporation; 2000.

Donders J. A confirmatory factor analysis of the California Verbal Learning Test-Second Edition (CVLT-II) in the standardization sample. Assessment. 2008; 15:123–131. [PubMed: 18187398]

Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. J Psychiat Res. 1975; 12:189–198. [PubMed: 1202204]

Gladsjo JA, Schuman CC, Evans JD, et al. Norms for letter and category fluency: demographic corrections for age, education, and ethnicity. Assessment. 1999; 6:147–178. [PubMed: 10335019]

Guillem F, Mograss M. Gender differences in memory processing: evidence from event-related potentials to faces. Brain Cogn. 2005; 57:84–92. [PubMed: 15629219]

Herlitz A, Nilsson LG, Bäckman L. Gender differences in episodic memory. Mem Cognit. 1997; 25:801–811.

Hermann BP, Wyler AR, Richey ET. Epilepsy, frontal lobe, and personality. Biol Psychiatry. 1987; 22:1055–1057. [PubMed: 3651526]

Hyde Z, Flicker L, Almeida OP, et al. Higher luteinizing hormone is associated with poor memory recall: the Health in Men Study. J Alzheimers Dis. 2010; 19:943–951. [PubMed: 20157250]

Jurica, PJ.; Leitten, CL.; Mattis, S. DRS-2 Dementia Rating Scale-2: Professional Manual. Lutz, FL: Psychological Assessment Resources; 2004.

Kramer JH, Delis DC, Kaplan E, et al. Developmental sex differences in verbal learning. Neuropsychology. 1997; 11:577–584. [PubMed: 9345701]

Kramer JH, Yaffe K, Lengenfelder J, et al. Age and gender interactions on verbal memory performance. J Int Neuropsychol Soc. 2003; 9:97–102. [PubMed: 12570363]

Lezak, MD.; Howieson, DB.; Loring, DW. Neuropsychological Assessment. New York: Oxford University Press; 2004.

Libon DJ, Bogdanoff B, Cloud BS, et al. Declarative and procedural learning, quantitative measures of the hippocampus, and subcortical white alterations in Alzheimer's disease and ischaemic vascular dementia. J Clin Exp Neuropsychol. 1998; 20:30–41. [PubMed: 9672817]

Libon DJ, Xie SX, Eppig J, et al. The heterogeneity of mild cognitive impairment: a neuropsychological analysis. J Int Neuropsychol Soc. 2010; 16:84–93. [PubMed: 19887015]

Mattis, S. Mental status examination for organic mental syndrome in the elderly patient. In: Bellak, L.; Karatsu, TB., editors. Geriatric Psychiatry. New York: Grune & Stratton; 1976. p. 77-121.

McGlone J. Sex differences in human brain asymmetry: a critical survey. Behav Brain Sci. 1980; 3:215–227.

Paolo AM, Troster AI, Ryan JJ. California Verbal Learning Test: normative data for the elderly. J Clin Exp Neuropsychol. 1997; 19:220–234. [PubMed: 9240482]

Phillips MD, Lowe MJ, Lurito JT, et al. Temporal lobe activation demonstrates sex-based differences during passive listening. Radiology. 2001; 220:202–207. [PubMed: 11425998]

Price CC, Garrett KD, Jefferson AL, et al. Leukoaraiosis severity and list-learning in dementia. Clinl Neuropsychol. 2009; 23:944–961.

Schmidt, M. Rey Auditory Verbal Learning Test: A Handbook. Los Angeles: Western Psychological Services; 1996.

Sims RC, Madhere S, Gordon S, et al. Relationships among blood pressure, triglycerides and verbal learning in African Americans. J Natl Med Assoc. 2008; 100:1193–1198. [PubMed: 18942281]

Stern Y, Blumen HM, Rich LW, et al. Space Fortress game training and executive control in older adults: a pilot intervention. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn. 2011; 18:653–677. [PubMed: 21988726]

Stricker JL, Brown GG, Wixted J, et al. New semantic and serial clustering indices for the California Verbal Learning Test-Second Edition: background, rationale, and formulae. J Int Neuropsychol Soc. 2002; 8:425–435. [PubMed: 11939700]

Wallentin M. Putative sex differences in verbal abilities and language cortex: a critical review. Brain Lang. 2009; 108:175–183. [PubMed: 18722007]

Wechsler, D. Wechsler Test of Adult Reading: WTAR. San Antonio, TX: The Psychological Corporation; 2001.

Woo E, Schmitter-Edgecombe M. Aging and semantic cueing during learning and retention of verbal episodic information. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn. 2009; 16:103–119. [PubMed: 18923945]

TABLE 1

Demographic and Neuropsychological Characteristics of the 2 Study Groups, by Sex

	California Verbal Learning Test-Second Edition				Philadelphia (repeatable) Verbal Learning Test			
	N	Women	Men	*P*	N	Women	Men	*P*
Age (years)	62	65.31 (4.19)	66.83 (3.97)	0.14	47	65.33 (8.42)	70.52 (15.78)	0.30
Education (years)	62	16.06 (1.90)	15.69 (3.11)	0.47	47	15.33 (2.19)	15.74 (2.37)	0.68
% Caucasian		63.90	65.40	0.90(χ^2)		82.10	68.40	0.27(χ^2)
Wechsler Test of Adult Reading Full-Scale Intelligence Quotient	59	107.84 (9.18)	111.03 (13.23)	0.10	42	104.52 (9.25)	110.46 (27.33)	0.54
Controlled Oral Word Association Test								
“CFL” form (average)	59	15.00 (4.05)	14.53 (4.44)	0.67	-	-	-	
“FAS” form (average)	-	-	-		44	17.13 (4.12)	16.46 (4.67)	0.61
Mattis Dementia Rating Scale (total)	62	139.81 (2.64)	140.78 (2.79)	0.17	-	-	-	
Mini-Mental State Examination (total)	-	-	-		47	29.24 (1.36)	29.19 (1.22)	0.87

NIH-PA Author Manuscript
NIH-PA Author Manuscript

TABLE 2

Results by Sex for the California Verbal Learning Test-Second Edition (CVLT-II) and Philadelphia (Repeatable) Verbal Learning Test (PrVLT)

	Total Immediate Recall		Semantic Clustering		Serial Clustering	
	CVLT-II	**PrVLT**	**CVLT-II**	**PrVLT**	**CVLT-II**	**PrVLT**
Women, mean (standard deviation)	50.92 (8.81)	39.00 (3.39)	1.08 (1.86)	1.27 (1.30)	0.67 (1.17)	0.08 (0.90)
Men, mean (standard deviation)	40.81 (7.50)	36.29 (4.45)	−0.02 (0.53)	0.46 (1.49)	0.77 (0.72)	0.35 (0.98)

TABLE 3

Summary of Regression Analyses for Model 1 with 2 Predictor Variables and Model 2 with 3 Predictor Variables, for the California Verbal Learning Test-Second Edition (CVLT-II) and Philadelphia (Repeatable) Verbal Learning Test (PrVLT)

	Total Recall		Semantic Clustering		Serial Clustering		*b*		*B*		sr^2_{unique}	
	CVLT-II	PrVLT	CVLT-II	PrVLT	CVLT-II	PrVLT	CVLT-II	PrVLT	CVLT-II	PrVLT	CVLT-II	PrVLT
Model 1												
Semantic clustering	0.52*	0.58*					4.97	1.88	0.80	0.66	0.45	0.20
Serial clustering	0.06	−0.38**	−0.56*	−0.73*			4.85	0.46	0.51	0.11	0.18	0
Model 2												
Semantic clustering	0.52*	0.58*					4.06	1.66	0.65	0.58	0.25	0.14
Serial clustering	0.06	−0.38**	−0.56*	−0.73*			4.20	0.31	0.44	0.07	0.13	0
Sex	−0.52*	−0.33**	−0.35**	−0.28***	0.05	0.12	−6.04	−1.44	−0.31	−0.18	0.08	0.03

The multiple regression is significant at

* $P \leq 0.001$,

** $P \leq 0.01$, or

*** $P \leq 0.05$.