Working memory in the oldest-old: evidence from output serial position curves

Emily M. Elliott, Katie E. Cherry, Jennifer Silva Brown, Emily A. Smitherman, S. Michal Jazwinski, Qingzhao Yu & Julia Volaufova

Springer

Springer

Working memory in the oldest-old: evidence from output serial position curves

Emily M. Elliott · Katie E. Cherry ·
Jennifer Silva Brown · Emily A. Smitherman ·
S. Michal Jazwinski · Qingzhao Yu · Julia Volaufova

Abstract In the present study, we examined adult age differences in short-term and working memory performance in middle-aged (45–64 years), young–old (65–74 years), old–old (75–89 years), and oldest–old adults (90 years and over) in the Louisiana Healthy Aging Study. Previous research suggests that measures of working memory are more sensitive to age effects than are simple tests of short-term memory Bopp and Verhaeghen (Journal of Gerontology: Psychological Sciences 60:223–233, 2005), Myerson, Emery, White, and Hale, (Aging, Neuropsychology, and Cognition 10:20–27, 2003). To test this hypothesis, we examined output serial position curves of recall data from three span tasks: forward and backward digit span and size judgment span. Participants' recall patterns in the size judgment span task revealed that the two oldest groups of adults showed the largest decreases in recall performance across output serial positions, but did not differ significantly from each other. Correlation analyses indicated the strongest negative correlation with age occurred with the size judgment span task. Implications of these findings for understanding strategic processing abilities in late life are discussed.

Corporate Author List Meghan B. Allen, BS; Gloria Anderson, BS; Iina E. Antikainen, BS; Arturo M. Arce, MD; Jennifer Arceneaux, RN; Mark A. Batzer, PhD; Emily O. Boudreaux, PhD; Lauri Byerley, PhD; Pauline A. Callinan, PhD; Catherine M. Champagne PhD, RD; Hau Cheng, MS; Katie E. Cherry, PhD; Yu-wen Chiu, MPH; Liliana Cosenza, BS; M. Elaine Cress, PhD-Consultant; Jianliang Dai, PhD; James P. DeLany, PhD; Jenny Y. Denver, PhD; Andy Deutsch, PhD; Melissa J. deVeer, MS; Devon A. Dobrosielski, MS; Rebecca Ellis, PhD; Andrea Ermolao, MD; Marla J. Erwin, MA; Mark Erwin, MA; Jennifer Fabre, MPT; Elizabeth T. Fontham, PhD; Madlyn Frisard, PhD; Paula Geiselman, PhD; Lindsey Goodwin, BS; Jason Gtray, BS; Valentina Greco, PhD; Tiffany Hall; Michael Hamilton, MD; Karri S. Hawley, PhD; Scott W. Herke, PhD; Jennifer Hayden, MS; Kristi Hebert, BS; Fernanda Holton, MA; Hui-Chen Hsu, PhD; S. Michal Jazwinski, PhD; Darcy Johanssen, PhD; Lumie Kawasaki, MD; Sangkyu Kim, PhD; Beth G. Kimball, BS; Christina King-Rowley, MS; Miriam Konkel, MD; Lynn LaMotte, PhD; Kim Landry; Carl Lavie, MD-Consultant; Daniel LaVie, BS; Matthew Leblanc; Christina M. Lefante, MPH; Li Li, MD; Hui-Yi Lin, PhD, MSPH; Kay Lopez, DSN; James Major; Bridget McEvoy-Hein, MS; John D. Mountz, MD PhD; Leann Myers, PhD; Jennifer Owens, BA; Kim B. Pedersen, PhD; Andrew Pellett, PhD; Eric Ravussin, PhD; Paul Remedios, BS; Ryan Rhodes, BS; Yolanda Robertson, NP; Jennifer Rood, PhD; Henry Rothschild, MD, PhD; Ryan A. Russell, BS; Erin Sandifer, BS; Charles Sasser, BS; Beth Schmidt, MS; Robert Schwartz, MD – Consultant; Donald K. Scott, PhD; Mandy Shipp, RD; Jennifer Silva Brown, PhD; F. Nicole Standberry; L. Joseph Su, PhD MPH; Jolie Thibodeaux; Mary Susan Thomas, APRN, MSN, FNP; Jessica Thomson, PhD; Valerie Toups, LPN; Crystal Traylor, APRN, MSN, WHNP; Cruz Velasco-Gonzalez, PhD; Julia Volaufova, PhD; Celeste Waguespack, BSN, RN; Jerilyn A. Walker, MS; Xui Yun Wang; David A. Welsh, MD; Michael A. Welsch, PhD; Robert H. Wood, PhD; Sarah Zehr, PhD; Pili Zhang, PhD. (Louisiana State University, Baton Rouge; Pennington Biomedical Research Center, Baton Rouge; Louisiana State University Health Sciences Center, New Orleans; Tulane University, New Orleans; University of Alabama, Birmingham).

E. M. Elliott · K. E. Cherry · J. S. Brown · E. A. Smitherman
Louisiana State University,
Baton Rouge, LA, USA

S. M. Jazwinski
Tulane University School of Medicine,
New Orleans, LA, USA

Q. Yu · J. Volaufova
Louisiana State University Health Science Center for the
Louisiana Healthy Aging Study,
New Orleans, LA, USA

E. M. Elliott (✉)
Department of Psychology, Louisiana State University,
Baton Rouge, LA 70803–5501, USA
e-mail: eelliott@lsu.edu

In the year 2009, the number of people in the United States aged 65 or older accounted for 13% of the population (or 39 million people). This number is projected to grow to 20% of the population by the year 2050 (or 88.5 million; U.S. Census, 2009). Recent demographic trends also indicate major increases in the segment of the population whose age falls into the category of oldest–old, yet research into the cognitive capabilities of these individuals is limited (Bäckman, Small, Wahlin, & Larsson, 2000). Focusing research efforts on the oldest segment of the adult population is therefore an important challenge for psychology.

In recent years, there has been increasing interest in the role of working memory (WM) as a critical component of adult cognition (Babcock & Salthouse, 1990; Bopp & Verhaeghen, 2005). WM is considered an active portion of the memory system; it is "a temporary storage system under attentional control that underpins our capacity for complex thought" (Baddeley, 2007, p. 1). A number of important cognitive functions have been linked to WM processes. For example, the link between WM and reading comprehension is well-established (Daneman & Carpenter, 1980), as is the relationship of WM and general fluid intelligence (Gray, Chabris, & Braver, 2003; Salthouse & Pink, 2008). Given the importance of WM to cognitive functioning, age-related differences have been investigated to improve our understanding of the causes of decline in WM performance (Craik & Bialystock, 2006; Dobbs & Rule, 1989). In the present study, we continued in this line of research by conducting an in-depth investigation of adult age differences in three measures of memory span, and extended previous research by including a large number of participants who fall into the age category of oldest–old (90 + years of age).

Researchers investigating the causes of age-related differences in WM have proposed several views, including single mechanism views (such as processing speed declines, e.g., Salthouse, 1996) and multiple mechanism views (such as dynamic contextual control and output forgetting, e.g. Maylor, Vousden, & Brown, 1999). The work of Maylor et al. was interpreted within the broader framework of a frontal-function hypothesis, given the large role of the frontal cortex in controlling behavior and its links to WM performance (Engle, 2001). Older adults may be more likely to suffer from changes in functioning of the frontal lobes, which leads to decreased executive functioning abilities (Myerson, Emery, White, & Hale, 2003). This change may also be interpreted in terms of the increased cognitive load that results from tasks requiring the coordination of processing and storage (see Barrouillet,

Bernardin, & Camos, 2004, for a discussion of cognitive load).

One matter of theoretical and practical importance concerns the assessment of the constructs of short-term and working memory. In particular, measurement practices have varied, with some researchers using simple span tasks requiring only repetition of items, such as the forward digit span task, and others using more complex span tasks requiring additional processing along with repetition, such as the backward digit span task (see Unsworth & Engle, 2007, for a discussion). Recent research suggests that both forward and backward digit span tasks can be used to effectively assess WM (Shelton, Elliott, Hill, Calamia, & Gouvier, 2009) in a sample of college-aged adults; however, the use of simple and complex span tasks with older adults, especially those older than 90 years of age, has not been undertaken to date.

Additionally, the use of serial position information has been central in the development of current theoretical views of working memory (Baddeley, 2007). Researchers examine the performance on each individual item within a list in a memory span task to look for trademark patterns of primacy (improved performance at the beginning of a list) and recency (improved performance at the end of a list). Examining serial position curves in recall can provide additional information about strategic processes that participants may engage in during a task. Very little research exists examining output serial order information in older adults (e.g., Golomb, Peelle, Addis, Kahana, & Wingfield, 2008; Maylor et al., 1999), and even less research exists investigating this important component of memory functioning in the oldest–old. In the Golomb et al. study, younger and older adults' free and serial recall of uncategorized word lists was examined. They found that older adults (M = 73 years) showed greater declines in serial recall performance than the younger adults (M = 20 years). Follow-up analyses examining the positions of the recalled items, relative to other items in the lists, revealed that older adults were less able to utilize order information. One specific example of this was that the contiguity effect differed by age; in free recall, younger adults tended to recall items together when these items had been originally presented together. Golomb et al.'s detailed analyses of order demonstrated that older adults have greater difficulty utilizing temporal context information and, furthermore, that older adults may try to rely on semantic information even when it is not helpful. However, the older adults were younger than the oldest group in the present sample, and further investigation of these issues with an older group is needed to reveal new information about the use of order information in very late life.

The primary objective of the present research was to provide new evidence on the strategic processing abilities

of very old adults. In particular, we examine hypotheses regarding age sensitivity in three different span tasks; we used both forward and backward digit span tasks, and the size judgment span task (SJS; Cherry, Elliott, & Reese, 2007) to evaluate WM performance. In the SJS task, participants heard items spoken aloud, and were asked to recall them in order of physical size, from smallest to largest. The items included were everyday concrete objects and animal names that were chosen for the ease with which a visual image could be created, thus allowing participants of varying educational and socioeconomic backgrounds to participate in the task (see also Cherry & Park, 1993). Important to consider within the SJS task was that, different from many WM measures that include both storage and processing concurrently, the processing demands most likely occurred after the items were been presented. Although it was possible that participants could try to encode new items while concurrently beginning the process of reordering them, this process would be increasingly difficult to manage as list length increased. Thus, particularly for longer list lengths, it seemed reasonable to assume that participants encoded the list of items and then engaged in a process of reordering them prior to beginning recall. This process is similar to other tasks involving reordering of items, such as backward digit span or letter–number sequencing tasks that are commonly included in the Wechsler Adult Intelligence Scale versions of assessment (Wechsler, 1997; see also Myerson et al., 2003). This property of the SJS task may place it at an intermediate level of difficulty between simple span tasks and more complex span tasks involving concurrent storage and processing demands (see the meta-analysis by Bopp & Verhaeghen, 2005), which makes it appropriate for use with the oldest–old segment of the population. Furthermore, this reordering of the items may be different from other types of tasks that involve separate item and order information, in which deficits in temporal context have been found with increased age (Hartman & Warren, 2005; Maylor et al., 1999).

The present focus on output serial position effects within both simple repetition and reordering tasks appears to be the first of its kind, especially with the inclusion of a large number of individuals 90 years of age and older. Given the known changes in WM performance over the lifespan, we hypothesized that significant age group differences would emerge, with the largest decreases in performance occurring in the oldest–old group. On the assumption that all three span measures tap into the global construct of WM (see Shelton et al., 2009; Unsworth, Spillers, & Brewer, 2010), this prediction should hold true across the three span measures included here. However, if forward digit span is less resource demanding than the other two span measures (i.e., backward digit span and SJS), then one might expect

to observe no significant age differences in the forward digit span task, but age group differences in the other two measures. Alternatively, previous research indicates that older adults have shown declines in temporal context information, and that this type of information may be more sensitive to age than are other types of associative memory (Hartman & Warren, 2005). Furthermore, research using the n-back paradigm demonstrated that these declines in temporal context affected performance on the n-back task to a greater degree than probable declines in item information (as assessed through item recognition and responses to lures; McCabe & Hartman, 2008). As mentioned earlier, the demands of the SJS task required that a participant be able to retain item information and then separate this information from the original order information to be able to report the items back in terms of their physical size. Thus, it is possible that older adults would rely upon item information and semantic knowledge, which do not show the same age-related declines as temporal order information, when performing the SJS task. In this case, the differences across the age groups may be less pronounced than in other types of WM tasks (such as complex span tasks; Lustig, May, & Hasher, 2001; McCabe & Hartman, 2003).

In the present research, we investigated adult age differences in memory for temporal order using three tasks, which varied in level of difficulty. The forward digit span task is a simple repetition task and relies heavily upon temporal order information. The backward digit span task also uses a small, closed set of items, and additionally requires participants to reorder these items. Both of these tasks tap temporal order information without a large contribution from semantic knowledge. If the oldest adults rely upon the combination of semantic and temporal order information more heavily than their younger counterparts, we may find poorer performance by the oldest–old adults in the forward and backward digit span tasks, relative to the SJS task, which includes semantic information. Vocabulary level is a factor to consider when discussing the role of semantic knowledge in the SJS task. Vocabulary level, often treated as a proxy for verbal intelligence, might be expected to play a larger role in the SJS task than for the tasks requiring the recall of simple digits. However, vocabulary knowledge does not show large declines across the lifespan (Cherry et al., 2008). We included a measure of vocabulary in the present study to assess the role of verbal intelligence in the performance of the span tasks.

Our second objective was to examine recall performance on the span tasks using multiple methods of scoring that account for overall levels of performance and output serial position performance. We expected that overall performance levels would differ among the age groups, as seen in previous research using span measures and the SJS (Bopp

& Verhaeghen, 2005; Cherry & Park, 1993; Cherry et al., 2008); however, the interaction of type of span measure and age group was of specific interest. Furthermore, detailed analyses were conducted using output serial position in these recall tasks, and age-related differences were evaluated using output serial position as the dependent measure (see also Maylor et al., 1999). On the basis of prior research, we expected to observe a significant interaction of age group with output serial position within each span measure. The comparison of the three types of span tasks in this population provided an opportunity to examine the hypothesis that larger effects of aging would emerge in the more complex tasks, relative to the simplest task (the forward digit span task; see Bopp & Verhaeghen, 2005; Myerson et al., 2003).

Method

Participants

Participants were drawn from the Louisiana Healthy Aging Study (LHAS), an ongoing multidisciplinary study of the determinants of longevity and healthy aging conducted in collaboration with Louisiana State University (LSU) Agricultural and Mechanical College, LSU Health Sciences Center in New Orleans, Tulane University, the University of Alabama at Birmingham, and the Pennington Biomedical Research Center in Baton Rouge, LA. LHAS participants are sampled randomly from the Voters Registration 2000 files for those age 20 to 64 years old, and from the Medicare Beneficiary Enrollment Data file of the Center for Medicare and Medicaid Services (CMS) for those age 65 years old and above for the eight parishes (counties) constituting the Greater Baton Rouge

community. Participants included 48 middle aged (45–59 years), 40 young–old (60–74 years), 40 old–old (75–89 years), and 90 of the oldest–old (90+ years), for a total of 218 participants. All scored at least a 25 on the Mini-Mental State Exam (MMSE; Folstein, Folstein, & McHugh, 1975) and were free of neurological impairment resulting from stroke or other trauma to the brain at the time of testing (see Table 1). Although participants completed a battery of tests, the present focus was on the forward and backward digit span tasks and the SJS task (see Cherry et al., 2008, for more details on the procedure). Participants in this study were compared on level of education, using a collapsed version of educational attainment that reduced the categories to two: either high school or above, or less than high school. A χ^2 test of independence revealed no significant association between educational attainment and age group, $\chi^2_{(3)} = 3.69$, $p = .29$.

Procedure

Participants were administered a number of different measures across two sessions. The measures of interest for the present investigation included a short form of the Wechsler Adult Intelligence Scale Vocabulary subtest (Jastak & Jastak, 1964) and the MMSE (Folstein et al., 1975) for a global indication of cognitive functioning, as well as the primary measures of interest, the forward and backward digit span tasks and the SJS task. Administration for all three tasks was stopped once a participant did not answer any sequences correctly at a given list length, or if the participant requested that the task be stopped.

During the forward digit span task, participants were read aloud lists of digits, and were asked to repeat them back in the same order that they were originally presented.

Table 1 Descriptive statistics for all participants

	M	Range	SD	Skew	Kurtosis	Reliability
Age	76.29	45-96	16.07	-.55	-1.25	–
MMSE	28.11	25-30	1.49	-.37	- .83	.98
Vocabulary	23.99	6 – 39	7.74	-.39	- .57	.96
SJS Score	3.96	2.00-6.50	0.78	.38	.85	.82
SJS Proportion 2-6	0.65	0.11-1.00	0.10	-.28	.03	.91
BDS Score	4.12	2.00-7.00	0.94	.62	.23	.75
BDS Proportion 2-7	0.63	0.17-0.98	0.11	.19	-.46	.92
FDS Score	5.69	2.50-8.50	1.00	.10	-.16	.72
FDS Proportion 2-8	0.50	0.19-1.00	0.13	.42	-.61	.91

Our measure of vocabulary was taken from Jastak & Jastak (1964), with the split-half reliability estimate drawn from the original work. Reliability indicates Cronbach's alpha for all other variables. *MMSE* Mini-Mental State Exam (Folstein et al., 1975; with the test–retest reliability estimate drawn from the 28-day comparison in the original work). *SJS* size-judgment span task, *BDS* backward digit span task, *FDS* forward digit span task, *Score* the all-or-none index of performance on the span tasks, *Proportion* the proportion of correct items including list lengths 2–6 for the SJS task, 2–7 for the BDS task, and list lengths 2–8 for the FDS task. $N = 218$

Lists began with three digits, and two lists of each length were administered until the length of nine digits was reached. The backward digit span task was administered in a similar fashion, with the exception that participants were asked to repeat the items in backward order from the original presentation. Lists began at two items in length and continued until eight items, with two lists at each length. In the SJS task, participants were read a list of easily imagined objects and were asked to repeat them back in order of physical size, from smallest to largest. For example, in a list length of three items, a participant might hear, "frog, piano, hairpin," and the correct response would be "hairpin, frog, piano." The administration of the task began with a practice sequence, two items in length. The test sequences consisted of list lengths of two to eight, with three instances of each length.

Results

The significance level was set to .05 for all analyses. Participants were first compared using two separate one-way ANOVAs, to assess age-related differences in performance on the vocabulary and MMSE measures. There were no significant age differences on the vocabulary scores ($p = .09$), but there were significant differences on the MMSE, $F (3, 214) = 15.84$, $MSE = 1.84$ (see Table 2). This analysis was followed by a Tukey's HSD using the harmonic mean sample size to correct for unequal groups. The post hoc analysis revealed that the oldest–old group had significantly lower MMSE scores than the three other age groups, which did not differ. The oldest–old age group was examined more closely by dividing the participants into groups, on the basis of their total MMSE scores, and calculating the mean of the SJS score for each group. The means were very similar across the range of MMSE scores (see Appendix A), suggesting that even though the oldest–old age group had a lower overall mean MMSE relative to the other age groups, performance on the SJS task was not differentially affected across MMSE scores within this age group.

Scoring of the three span tasks was done in two steps, to provide global and specific indices of performance (see Table 1 for descriptive statistics). An overall span score,

referred to as the FDS Score, BDS Score, or SJS Score, was calculated by counting the number of list lengths in which perfect recall was achieved, with an additional half a point for list lengths in which at least one list was answered correctly (for example, all three lists at list lengths 2 and 3, plus one list length correct at list length 4 = span score of 3.5). The specific index of performance, referred to as either FDS, BDS, or SJS proportion, was calculated by giving credit for each individual item correctly recalled, regardless of whether or not the entire sequence was recalled correctly across the range of possible list lengths (2–9 for FDS; 2–8 for BDS and SJS). This method allowed for a proportion of correct recall to be calculated for each sequence, by dividing the number of correctly recalled items by the total number of items available in that sequence.

Methods of scoring: overall scores

The analyses began with an age group x span measure mixed-model ANOVA to examine the main effects and potential interaction of age group with type of span measure (see Fig. 1 for the means and standard errors). This analysis revealed significant main effects of age group, $F (3, 214) = 13.33$, $MSE = .04$, $\eta_p^2 = .16$, span measure, $F (2, 428) = 28.90$, $MSE = .014$, $\eta_p^2 = .12$, and a significant interaction of these two factors, $F (6, 428) = 2.75$, $MSE = .014$, $\eta_p^2 = .04$. To investigate the hypothesis that the age groups will have differential patterns of performance on the three span measures, the significant interaction was followed by one-way ANOVAs for each task separately.

A one-way ANOVA on the overall SJS score yielded a significant age group effect, $F (3, 214) = 28.78$, $MSE = .45$. A Tukey's HSD using the harmonic mean sample size to

Table 2 Mean total MMSE and vocabulary scores (and standard deviations) by age group

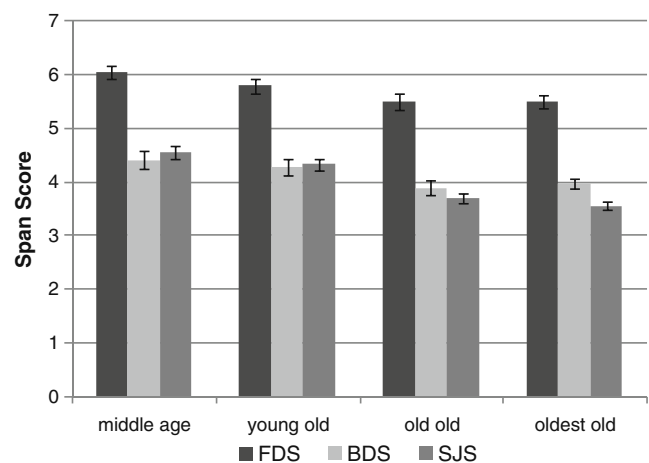| Age Group	MMSE	Vocabulary	N
Middle age	28.79 (1.37)	25.33 (7.78)	48
Young–old	28.78 (1.21)	24.93 (7.79)	40
Old–old	28.23 (1.23)	25.00 (7.15)	40
Oldest–old	27.39 (1.46)	22.40 (7.80)	90

Fig. 1 This figure depicts the average span score for each of the three span measures for each of the four age groups. *Error bars* represent standard errors of the means. *FDS* forward digit span, *BDS* backward digit span, and *SJS* size judgment span

correct for unequal groups revealed that the middle-aged group and the young–old groups did not differ from each other, but that both differed from the old–old and oldest–old groups, which also did not differ from each other. For a detailed look at the differences, the frequency counts of each SJS score, by age group, are presented in Appendix B.

This same analysis was conducted on the overall BDS score, with a significant age group effect, $F (3, 214) = 3.31$, $MSE = .86$. The Tukey's HSD revealed that the means of the four age groups differed only for the two extremes, $M = 3.94$ for the old–old age group, and $M = 4.42$ for the young–old age group.

Finally, the FDS score was examined in a similar analysis, resulting in a significant age group effect, $F (3, 214) = 3.68$, $MSE = .96$. The Tukey's HSD post hoc analysis indicated that the youngest age group had a significantly higher mean score ($M = 6.06$) than the other age groups, which did not differ from each other.

Methods of scoring: proportions

Performance across the list lengths was examined using the proportion-correct scores as the dependent measure, for the different age groups. First, results from the SJS task will be discussed, followed by the BDS and FDS tasks. Because of floor effects at the longest list lengths, these highest list lengths were removed from the following analyses to make the comparison of the proportion-correct measures more comparable to the "score" measures. That is, these higher lists would not be included in the scores unless they were answered correctly; the proportion-correct measure can include one correct item within an otherwise completely incorrect list, thus "artificially" lowering the proportion-correct scores by including the highest list lengths.

Because of floor effects at list lengths of 7 and 8 in the SJS task, these lists were dropped from the remaining analyses. A mixed-model-type ANOVA was conducted on the average proportion-correct scores to adjust for the repeated measures on the same subject at different list lengths. The main effects are age group, list length, and we also include the Age Group x List Length interaction in the model. The analysis revealed significant main effects of age group, $F (3, 214) = 35.79$, $MSE = .08$, $\eta_p^2 = .32$, list length, $F (4, 856) = 889.76$, $MSE = .027$, $\eta_p^2 = .81$, and a significant interaction of these two factors, $F (12, 856) = 9.74$, $MSE = .027$, $\eta_p^2 = .12$. These significant findings can be described as a decrease in the mean performance scores across both increasing age and increasing list length, with a significant drop in proportion-correct scores for the two oldest age groups at the longest list lengths (see Fig. 2a).

Similar analyses were conducted for the BDS task, and performance was very poor on the highest list length (list length 8). The following analysis only included list lengths
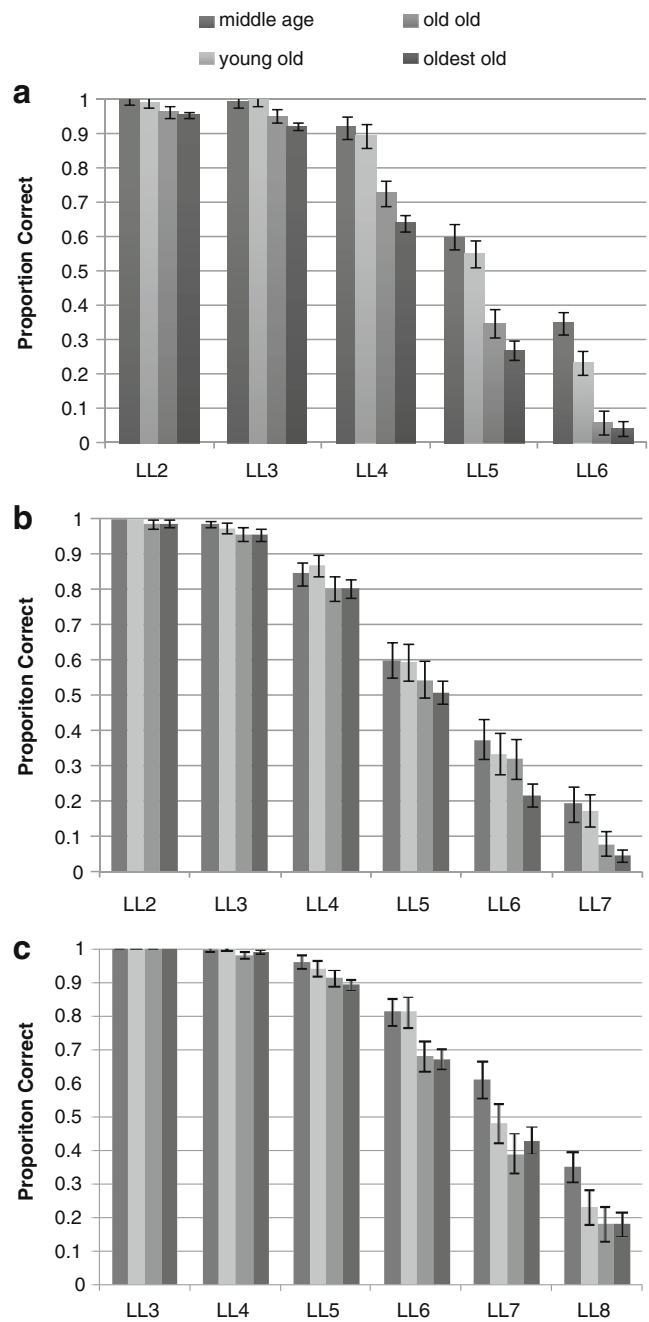

Fig. 2 a This figure depicts the average of the proportion-correct scores for list lengths 2–6 in the size-judgment span task for each age group. *Error bars* represent standard errors of the means. b This figure depicts the average of proportion-correct scores for list lengths 2–7 in the backward digit span task for each age group. *Error bars* represent standard errors of the means. c This figure depicts the average of proportion-correct scores for list lengths 2–8 in the forward digit span task for each age group. *Error bars* represent standard errors of the means

2–7 (see Fig. 2b). The mixed-model type ANOVA revealed significant main effects of age group, $F (3, 214) = 2.76$, $MSE = .16$, $\eta_p^2 = .04$, and list length, $F (5, 1070) = 645.35$, $MSE = .039$, $\eta_p^2 = .75$. The interaction of these two factors

was marginally significant, $p = .07$. The change in performance across the age groups in the BDS task was similar to that observed in the SJS task. The average BDS score decreased as the age and/or list length increased.

The FDS task also demonstrated floor effects at the longest list length. The present analysis includes only lists 2–8 (see Fig. 2c, which begins with list length 3 because of ceiling effects at list length 2). The results for the FDS task were very similar, revealing two significant main effects: list length, $F(6, 1284) = 994.47$, $MSE = .006$, $\eta_p^2 = .82$, and age group, $F(3, 214) = 4.80$, $MSE = .24$, $\eta_p^2 = .06$, with a significant interaction, $F(18, 1284) = 1.69$, $MSE = .006$, $\eta_p^2 = .02$. The reduction in overall level of difficulty of the FDS task is observed by noting that participants continued to perform well into higher list lengths as compared with performance in the two previous span measures.

Output serial position effects

To examine performance more closely in the SJS task, list length 4 was chosen, since it was not constrained by either ceiling or floor effects for the different age groups (see Fig. 2a). The average of performance on each item within list length 4 was the dependent measure in this analysis of output serial position effects (see Fig. 3a). An item was scored in the correct output serial position if it was reported in the correct order of physical size, from smallest to largest item. A mixed-model-type ANOVA accounting for the repeated measures of SJS over the same subject with the main effects of age group, serial position, and Age Group x Serial Position interaction was conducted, revealing significant main effects of age group, $F(3, 214) = 21.60$, $MSE = .21$, $\eta_p^2 = .23$, serial position, $F(3, 642) = 72.99$, $MSE = .03$, $\eta_p^2 = .25$, and a significant interaction of these two factors, $F(9, 642) = 6.03$, $MSE = .03$, $\eta_p^2 = .08$. This interaction was followed up with a series of repeated measures ANOVAs separately for each age group. Although each age group revealed a significant main effect of serial position [middle-age, $F(3, 141) = 5.31$, $MSE = .01$, $\eta_p^2 = .10$; young–old, $F(3, 117) = 11.54$, $MSE = .02$, $\eta_p^2 = .23$; old–old, $F(3, 117) = 28.71$, $MSE = .04$, $\eta_p^2 = .42$; oldest–old, $F(3, 267) = 49.59$, $MSE = .04$, $\eta_p^2 = .36$], follow-up

pairwise comparisons reveal that positions 1 and 2 differed significantly from positions 3 and 4 in the middle-age and young–old age groups. All positions differed from each
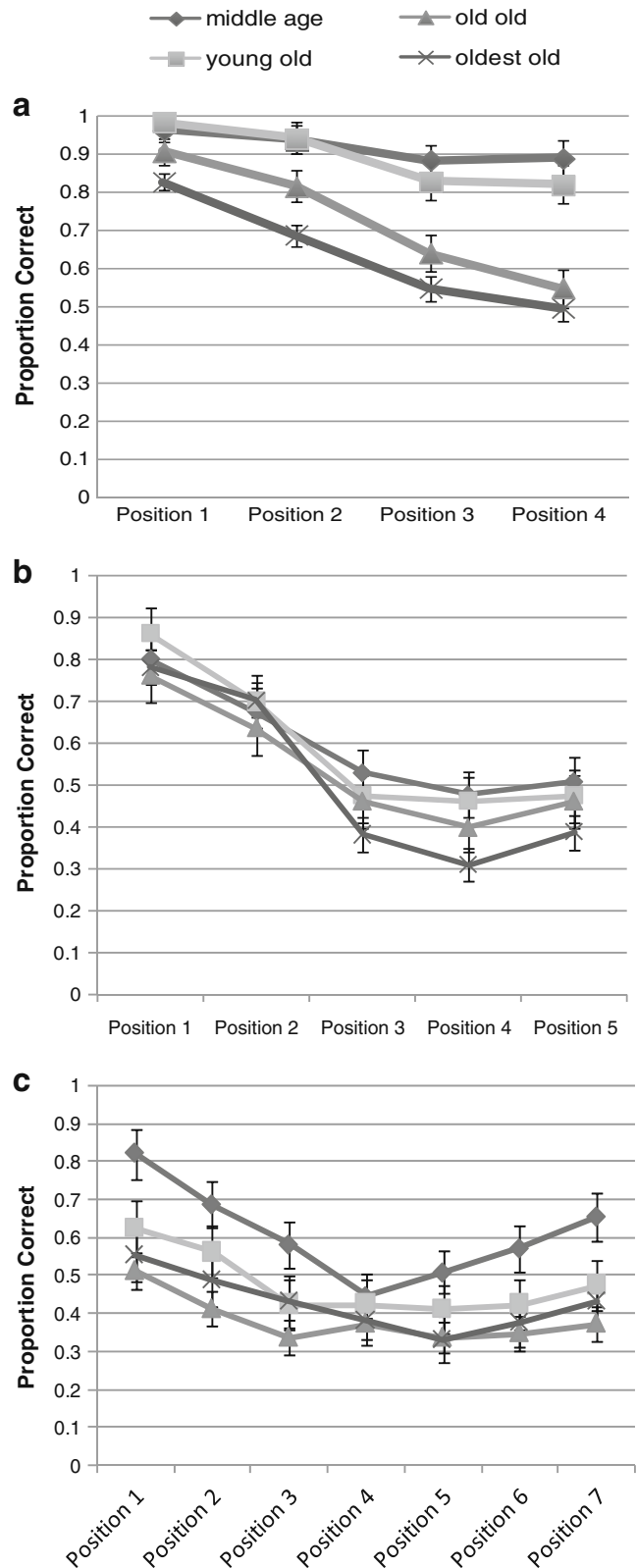

Fig. 3 a. This figure depicts the average of the proportion-correct scores for output serial positions 1–4 for list length 4 in the SJS task, for each age group. *Error bars* represent standard errors of the means. **b**. This figure depicts the average of the proportion-correct scores for output serial positions 1–5 for list length 5 in the backward digit span task, for each age group. *Error bars* represent standard errors of the means. **c**. This figure depicts the average of the proportion-correct scores for output serial positions 1–7 for list length 7 in the forward digit span task, for each age group. *Error bars* represent standard errors of the means

other for the old–old age group, and all pairings except for the positions 3 and 4 were different in the oldest–old age group. Thus, the pattern of output serial positions differed among the age groups, with the exception that the two youngest groups of participants displayed the same patterns in their output serial position curves.

The backward digit span task was analyzed at list length 5, again, to avoid ceiling or floor effects (see Fig. 2b). The serial position effects appear in Fig. 3b. A mixed-model-type ANOVA revealed a significant main effect of serial position, $F(4, 856) = 99.97$, $MSE = .06$, $\eta_p^2 = .32$, and a significant interaction of the factors of age group and serial position, $F(12, 856) = 1.87$, $MSE = .06$, $\eta_p^2 = .03$. However, the main effect of age group was not significant ($p = .41$). The significant interaction was followed up with a series of repeated measures ANOVAs separately for each age group. The results of these analyses were significant for each age group [middle-age, $F(4, 188) = 17.17$, $MSE = .05$, $\eta_p^2 = .27$; young–old, $F(4, 156) = 20.86$, $MSE = .06$, $\eta_p^2 = .35$; old–old, $F(4, 156) = 18.63$, $MSE = .05$, $\eta_p^2 = .32$; oldest–old, $F(4, 356) = 70.46$, $MSE = .06$, $\eta_p^2 = .44$]. Pairwise comparisons revealed that each age group showed a similar pattern across the serial positions, with positions 1 and 2 showing much higher levels of performance, and differing significantly from all positions, whereas positions 3, 4, and 5 did not differ from each other. The oldest–old showed the largest differences in performance at the later serial positions where only positions 3 and 5 are not significantly different.

The forward digit span task had overall higher levels of performance, and was thus examined at list length 7 (see Fig. 2c). The serial position effects for the forward digit span task are depicted in Fig. 3c. As before, a mixed-model-type ANOVA was conducted, revealing significant main effects of age group, $F(3, 214) = 3.30$, $MSE = .99$, $\eta_p^2 = .04$, serial position, $F(6, 1284) = 30.66$, $MSE = .04$, $\eta_p^2 = .13$, and a significant interaction of these two factors, $F(18, 1284) = 1.63$, $MSE = .04$, $\eta_p^2 = .02$. Repeated measures ANOVA were conducted separately to examine the interaction in more depth. For the present task, we asked participants to listen to a sequence of digits, and to repeat them back in forward order. Thus, one would expect to see traditional patterns of performance in recall from the auditory modality: primacy (better recall at the beginning of the list) and recency (better recall for the final item in the list), with poorer performance for the items in the middle of the list. Although it is easy to note the primacy effects (see Fig. 3c), the recency effects are less clearly observed, especially in the older age groups of participants, and the middle-aged group showed the most traditional serial position curve. This interpretation was supported by the findings from the ANOVAs and the pairwise comparisons. All four age groups demonstrated

significant changes across serial positions [middle-age, $F(6, 282) = 16.71$, $MSE = .04$, $\eta_p^2 = .26$; young–old, $F(6, 234) = 6.17$, $MSE = .05$, $\eta_p^2 = .14$; old–old, $F(6, 234) = 3.41$, $MSE = .05$, $\eta_p^2 = .08$; oldest–old, $F(6, 534) = 11.97$, $MSE = .04$, $\eta_p^2 = .12$. The pairwise comparisons indicated that all serial positions differed from the first position in the middle-aged group (demonstrating primacy); however, in the young–old group, only positions 3, 4, and 5 differed significantly from position 1. In contrast, the old–old age group revealed the flattest curve, with no significant differences emerging among the seven serial positions. Finally, the oldest–old revealed strong primacy effects, with all but position 2 differing significantly from position 1.

Correlation analyses

A comparison of the scores and the proportion measures (including the list lengths indicated previously for each task) was conducted to address the second objective of the study, which was to determine which of these methods allowed for a stronger relationship with other indicators of cognitive aging. For this analysis, correlations were conducted with the multiple methods of scoring SJS, BDS, and FDS; the MMSE; and vocabulary (see Table 3). MMSE and vocabulary scores were chosen for inclusion in the correlation analyses as indicators of global cognitive functioning and verbal intelligence, respectively. In the first analysis, age was included as a variable, and in the second, the correlations were conducted with age partialled out. In terms of the level of difficulty hypothesis of the three span tasks, the patterns of correlations suggested that the SJS, with similar outcomes for both scoring measures, was strongly and negatively correlated with age (rs = −.55 and −.58 for the SJS score and SJS proportion, respectively). The other two span measures, FDS and BDS, had correlations with age of similar magnitudes to each other, again regardless of scoring. In particular, the r values were −.23 and −.26 for FDS score and FDS proportion, respectively, whereas the r values, in order, were −.21 and −.20 for BDS score and BDS proportion, respectively. This pattern of correlations would argue against an increasing level of difficulty as one moved from the FDS task, to BDS. It instead suggests that the demands of the SJS task were more strongly related to differences with increased age, and the FDS and BDS shared similar relationships to each other with age.

These relationships were explored by examining the correlations of the span tasks with the MMSE and vocabulary scores as indicators of cognitive functioning. The MMSE total score was strongly correlated with age ($r = -.41$), as were the measures of SJS span (rs = .37 and .36 for the SJS score and SJS proportion, respectively). In contrast,

Table 3 Correlations among measures with age, and with age partialled out

	1	2	3	4	5	6	7	8	9
Age (1)									
MMSE (2)	-0.41 **		0.37 **	0.19 **	0.17 *	0.16 *	0.21 **	0.03	0.00
Vocabulary (3)	-0.15 *	0.40 **		0.26 **	0.27 **	0.14 *	0.28 **	0.28 **	0.24 **
SJS Score (4)	-0.55 **	0.37 **	0.30 **		0.84 **	0.35 **	0.38 **	0.41 **	0.42 **
SJS Proportion 2–6 (5)	-0.58 **	0.36 **	0.30 **	0.87 **		0.33 **	0.35 **	0.37 **	0.37 **
BDS Score (6)	-0.21 **	0.22 **	0.17 *	0.40 **	0.36 **		0.78 **	0.38 **	0.41 **
BDS Proportion 2-7(7)	-0.20 **	0.27 **	0.30 **	0.41 **	0.38 **	0.79 **		0.41 **	0.44 **
FDS Score (8)	-0.23 **	0.12	0.31 **	0.46 **	0.42 **	0.41 **	0.43 **		0.93 **
FDS Proportion 2– 8 (9)	-0.26 **	0.11	0.27 **	0.48 **	0.43 **	0.44 **	0.47 **	0.93 **	

Values above the diagonal, in bold print, represent the correlations with age partialled out.

$*p < .05$, $** p < .01$

FDS was not correlated with MMSE, suggesting that the requirements of these two tasks differed greatly. Furthermore, these correlations went to zero when age was partialled out. The BDS measures shared modest, but significant, correlations with the MMSE ($rs = .22$ and .27 for BDS score and BDS proportion, respectively). Finally, the relationship between vocabulary scores and the SJS task was considered to determine whether vocabulary as a proxy for verbal knowledge contributed significantly to overall performance on the SJS task. The correlation was positive and significant ($r = .30$), although this relationship was not heavily driven by contributions of vocabulary knowledge differentially for the different age groups, given that this relationship did not decrease much in strength when age was partialled out ($r_{partial} = .26$). The relationship of vocabulary scores and age was not very strong overall ($r = −.15$).

General discussion

Recent demographic trends reveal large increases in the percentages of the oldest members of our society. It is increasingly important to understand the cognitive functioning of older adults, with special emphasis placed on the oldest–old segment (those 90+years of age). In the present study, we addressed this concern by conducting an in-depth analysis of performance on three measures of memory in a sample of participants ranging in age from 45 to 96 years of age, with a total of 90 participants in the 90+age group. We were interested in potential differences in strategic processing across the age groups, as well as in the sensitivity of three measures of memory span to reveal such differences.

We compared performance on three measures of memory span using both an all-or-none scoring method and a method that allowed for proportion-correct scoring. Both scoring methods revealed that adults had increasing difficulty with these tasks as age increased (see Figs. 2a–c). Importantly, output serial position data were examined to determine how the demands of reordering the items in terms of their physical size would affect participants' performance, relative to the output serial position curves of the two digit span tasks. The size judgment span task placed additional demands upon the particpants, beyond simply repeating the series of items. To perform this task correctly, participants had to maintain the items during the encoding phase, and then reorder them prior to recall. The data suggested that the oldest participants had the most difficulty with the size judgment span task, as revealed by the overall scores, output serial position curves, and the significant, negative correlation of the size judgment span task with age. Thus, the size judgment span task revealed the most sensitivity to age-group differences. Our results are similar to Myerson et al.'s (2003) findings in which the largest declines with increasing age were observed in the letter–number sequencing task, as compared with forward and backward digit span tasks. The letter–number sequencing task places similar demands on participants as does the size judgment task, in that participants must categorize the letters and numbers within the list of items separately and report them in their respective orders (alphabetical and numeric).

The individual tasks were then examined in detail, using the output serial position curves as an index of strategic processing. Typical serial position curves were not expected in the size judgment span task, because it invovled reordering. All age groups revealed primacy effects in the output serial position data for list length 4, and no recency. Primacy effects typically relate to rehearsal of the items; however, in this reordering task, the relationship between input and output serial position

is not always maintained. The high performance levels at output serial position 1 indicate that the smallest item was tagged within the list, but additional ordering by size became more difficult as the list continued, differentially, for the oldest adults.

The two younger age groups had serial position curves that remained relatively flat, whereas the two older age groups had much larger drops in performance across the list items. The marked decreases in memory of the later list items was much greater in the older two groups of participants, suggesting that the requirements of not only retaining the items, but then placing them in a specific order, was much more difficult for these older participants. This pattern was reflected in the post hoc comparisons of output serial position performance within each age group. Interestingly, this post hoc analysis did not show significant differences in output serial position performance between the two oldest age groups, suggesting that these participants may have been able to draw upon item information and semantic associations. The use of semantic associations has shown less steep declines with age, relative to item information (see Hartman & Warren, 2005; McCabe & Hartman, 2008). Future research with the two oldest age groups should continue to investigate the trajectory of the temporal context deficit. The work of Maylor et al. (1999) supported the hypothesis of context deficits, but did not include adults in the 90+ age range. Comparing the pattern of performance in the size judgment span task to the backward digit span task illustrated the differences in difficulty level of these two tasks, since a higher list length could be examined without the problems of floor effects. Although both tasks required participants to do reordering of the items, using familiar digits led to higher performance than did ordering items by their physical size. Similar findings of primacy effects emerged as with the size-judgment span task, but the sharp differences among the four age groups were not observed. Although the recall performance of the oldest–old decreased more in the later serial positions, the pairwise comparisons for the four age groups did not result in different patterns. All four age groups showed primacy effects: positions 1 and 2 differed significantly from positions 3, 4, and 5.

Finally, the forward digit span task was investigated, and results revealed that this task was easier than the other two, in terms of higher levels of performance at higher list lengths. List length 7 was examined in detail, and again, strong primacy effects were found in all age groups. Although the general bow-shaped pattern of serial recall was evident in all groups, it appeared most clearly in the middle-aged group. The shape of the curve suggests that the middle-aged group is relying more heavily upon rehearsal processes, whereas the older adults may be employing multiple strategies to perform the task successfully. The oldest–old did not have the lowest levels of performance across all of the list lengths in this task. The role of experience with familiar digits, and the lower level of difficulty of the forward digit span task, may have allowed for the use of multiple strategies without major decrements in performance across the age groups.

Taken together across the three span measures, the pattern of performance of the oldest–old was not dramatically different than that of their younger counterparts. In many analyses, they performed very similarly to the old–old age group. However, one methodological limitation of the present research warrants brief mention. Those who participate in a multidisciplinary study of healthy aging may be unusually high functioning. Conceivably, our results may reflect selectivity in the direction of cognitive vitality in late life, which necessarily limits the generalizability of these results. Notwithstanding the health of our oldest–old, the MMSE scores were significantly lower in this age group, reflecting the presence of healthy aging (see Table 2 and Appendix A). Future studies will continue to address the issues that relate to longevity, and the potential relationship of longevity to cognitive functioning.

In conclusion, these data provided a unique window into the strategic processing abilities of the oldest–old in three memory tasks. Examining output serial position effects allowed for a more detailed understanding of performance and revealed changes across serial positions for the different age groups. The pattern of results suggested that the two younger age groups performed similarly, and that the two older age groups performed similarly. These data confirmed previous findings that working memory abilities decrease with increasing age, and supported the view that changes in the use of semantic knowledge may be less sensitive to changes with age (see Hartman & Warren, 2005; McCabe & Hartman, 2008). Furthermore, the use of complex measures, such as the size judgment span task, which required both memory for items and a reordering of these items, proved to be the most sensitive to age differences, whereas the forward digit span task was less sensitive. Our results replicated similar findings in the literature when forward and backward digit span tasks have been compared across increasing ages (Bopp & Verhaeghen, 2005; Gregoire & Van der Linden, 1997; Myerson et al., 2003), and provided support for the view of aging that includes frontal-function deficits, or changes in executive control (Maylor et al., 1999; Myerson et al., 2003). In addition, these patterns were shown to exist in the oldest–old.

Appendix A

Table 4 Mean SJS scores in the oldest–old as a function of the MMSE total score

MMSE total	*M* of SJS Score	*N*
25	3.55	10
26	3.31	18
27	3.60	19
28	3.68	19
29	3.64	18
30	3.75	6

Appendix B

Table 5 Means and frequency counts of SJS score by age group

	Middle Age	Young–Old	Old–Old	Oldest–Old
2.0	–	–	–	4
2.5	–	–	–	3
3.0	1	1	12	22
3.5	3	4	6	18
4.0	16	18	15	35
4.5	11	7	5	6
5.0	9	6	2	2
5.5	4	2	–	–
6.0	2	2	–	–
6.5	2	–	–	–
Total	48	40	40	90
Mean	4.56	4.33	3.74	3.57

References

Babcock, R. L., & Salthouse, T. A. (1990). Effects of increased processing demands on age differences in working memory. *Psychology and Aging, 5,* 421–428.

Bäckman, L., Small, B. J., Wahlin, A., & Larsson, M. (2000). Cognitive functioning in very old age. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 499–558). Mahwah: Erlbaum.

Baddeley, A. (2007). *Working memory, thought, and action.* New York: Oxford University Press.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resources sharing in adults' working memory spans. *Journal of Experimental Psychology. General, 133,* 83–100.

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *Journal of Gerontology: Psychological Sciences, 60,* 223–233.

Cherry, K. E., Elliott, E. M., & Reese, C. M. (2007). Age and individual differences in working memory: The size judgment span task. *The Journal of General Psychology, 134,* 43–65.

Cherry, K. E., Hawley, K. S., Jackson, E. M., Volaufova, J., Su, L. J., & Jazwinski, S. M. (2008). Pictorial superiority effects in oldest-old adults. *Memory, 16,* 728–741.

Cherry, K. E., & Park, D. C. (1993). Individual difference and contextual variables influence spatial memory in younger and older adults. *Psychology and Aging, 8,* 517–526.

Craik, F. I., & Bialystock, E. (2006). Cognition through the lifespan: Mechanisms of change. *Trends in Cognitive Sciences, 10,* 131–138.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Behavior, 19,* 450–466.

Dobbs, A. R., & Rule, B. G. (1989). Adult age differences in working memory. *Psychology and Aging, 4,* 500–503.

Engle, R. W. (2001). What is working memory capacity? In H. L. Roediger, H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 297–314). Washington, DC: American Psychological Association.

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

Golomb, J. D., Peelle, J. E., Addis, K. M., Kahana, M. J., & Wingfield, A. (2008). Effects of adult aging on utilization of temporal and semantic associations during free and serial recall. *Memory & Cognition, 36,* 947–956.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Neuroscience, 6,* 316–322.

Gregoire, J., & Van der Linden, M. (1997). Effects of age on forward and backward digit spans. *Aging, Neuropsychology, and Cognition, 4,* 140–149.

Hartman, M., & Warren, L. H. (2005). Explaining age differences in temporal working memory. *Psychology and Aging, 20,* 645–656.

Jastak, J., & Jastak, S. (1964). Short forms of the WAIS and WISC vocabulary subtests. *Journal of Clinical Psychology, 20,* 167–199.

Lustig, C., May, C. P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology. General, 130,* 199–207.

Maylor, E. A., Vousden, J. I., & Brown, G. D. A. (1999). Adult age differences in short-term memory for serial order: Data and a model. *Psychology and Aging, 14,* 572–594.

McCabe, J., & Hartman, M. (2003). Examining the locus of age effects on complex span tasks. *Psychology and Aging, 18,* 562–572.

McCabe, J., & Hartman, M. (2008). Working memory for item and temporal information in younger and older adults. *Aging, Neuropsychology, and Cognition, 15,* 574–600.

Myerson, J., Emery, L., White, D. A., & Hale, S. (2003). Effects of age, domain, and processing demands on memory span: Evidence for differential decline. *Aging, Neuropsychology, and Cognition, 10,* 20–27.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403–428.

Salthouse, T. A., & Pink, J. E. (2008). Why is working memory related to fluid intelligence? *Psychonomic Bulletin & Review, 15,* 364–371.

Shelton, J. T., Elliott, E. M., Hill, B. D., Calamia, M. R., & Gouvier, W. D. (2009). A comparison of laboratory and clinical working memory tests and their prediction of fluid intelligence. *Intelligence, 37,* 283–293.

United States Census (2009). *Older Americans month: May 2009.* Retrieved March 26, 2009 from http://www.census.gov/Press-Release/www/releases/archives/facts_for_features_special_editions/013384.html

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin, 133,* 1038–1066.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2010). The contributions of primary and secondary memory to working memory capacity: An individual differences analysis of immediate free recall. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 36,* 240–247.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-Third Edition: Technical Manual*. San Antonio: The Psychological Corporation.

Author note