Published in final edited form as:
Memory. 2019 October ; 27(9): 1236–1249. doi:10.1080/09658211.2019.1645859.

Age-Related Differences in Metacognition for Memory Capacity and Selectivity

Alexander L. M. Siegel, **Alan D. Castel**
Department of Psychology, University of California, Los Angeles.

Abstract

Background—We compared two types of metacognitive monitoring in younger and older adults: metacognitive accuracy for their overall memory performance and their ability to selectively remember high-value information.

Method—Participants studied words paired with point values and were asked to maximize their point score. In Experiment 1, they predicted how many words they would remember while in Experiment 2, they predicted how many points they would earn.

Results—In Experiment 1, while younger adults were accurate in their predictions, older adults were overconfident in the number of words they would recall throughout the task. In Experiment 2, however, both younger and older adults were equally accurate when predicting the amount of points they would earn after some task experience.

Conclusions—While younger adults may have higher metacognitive accuracy for their capacity, older adults can accurately assess their ability to selectively remember information, suggesting potentially separate metacognitive mechanisms that are differentially affected by aging.

Keywords

memory; metacognition; predictions; selectivity; value-directed remembering

---

Metacognition, the ability to monitor and control our cognitive processes, is a crucial aspect of daily functioning. Metamemory, the metacognitive processes associated with memory, allows us to assess memory quality or strength and adjust our behavior to regulate our memories. For example, when learning information for an upcoming exam, it is imperative for a successful student to accurately evaluate their knowledge of the material (e.g., "How well do I know this piece of information?") and adjust their behavior to account for this evaluation (e.g., "I do not know it that well, so I need to study this information in more depth"). Metacognitive functioning is also critical in old age when memory errors may be more frequent. For example, older adults must remember which medications they have taken in a given day and must be able to adjust their behavior in order to account for this assessment (e.g., "I forgot to take my blood pressure medication earlier, so I must do so now"). As such, it is important for younger and older adults to accurately monitor their

Correspondence concerning this article should be directed to Alex Siegel, 1285 Psychology Tower, Department of Psychology, University of California, Los Angeles, Los Angeles, CA 90095. alexsiegel93@ucla.edu.

memory performance and subsequently control their behaviors to maximize this performance

Effective metacognitive functioning may become more important as we age due to an increase in the frequency of episodic memory errors (Hertzog & Dixon, 1994). Thus, the ability to monitor when information will be later remembered or forgotten may be a particularly important skill for older adults. In contrast to well-documented episodic memory deficits that occur with advancing age (for a review, see Hess, 2005; Zacks & Hasher, 2006), metacognitive processes associated with memory may experience little to no age-related decline in some circumstances (Castel, Middlebrooks, & McGillivray, 2016; Hertzog & Dunlosky, 2011). Various metamemory studies utilizing judgments of learning (JOLs) to examine how well participants can assess whether information will be later recalled have found negligible differences in JOL accuracy between younger and older adults (Hertzog, Sinclair, & Dunlosky, 2010; Hines, Touron, & Hertzog, 2009). Additional work has shown that older adults are equally as accurate as younger adults in determining when and how much information they may have forgotten between initial encoding and retrieval (Halamish, McGillivray, & Castel, 2011)

Importantly, this lack of age-related differences in JOL accuracy may only be the case when judgments are made on a local, item-by-item basis. Other work has demonstrated that, when asked to make global predictions about recall performance on an entire set of to-be-remembered materials, age-related differences are observed, as older adults may be overconfident in their memory performance (Bruce, Coyne, & Botwinick, 1982; Connor, Dunlosky, & Hertzog, 1997; Hertzog, Saylor, Fleece, & Dixon, 1994; cf. Kavé & Halamish, 2015). Older adults may also be overconfident in predicting how much information will be accompanied by recollective experience (as compared to feelings of familiarity or knowing), suggesting that there are also age-related declines in the monitoring of recollection (Soderstrom, McCabe, & Rhodes, 2012). Thus, while older adults' item-by-item metacognitive processing may be relatively unimpaired by aging, the application of the information gained from this monitoring to make a global assessment may be difficult for older adults

One instance in which older adults demonstrate successful metacognitive functioning is illustrated by their memory performance on a value-directed remembering (VDR) task (Ariel, Price, & Hertzog, 2015; Castel, Benjamin, Craik, & Watkins, 2002; Castel, 2008; Hayes, Kelly, & Smith, 2012). In this type of task, younger and older adult participants are presented a series of words paired with point values indicating their importance. Participants study each word for a set amount of time with the goal of maximizing their point score (a summation of the points associated with correctly recalled words). Results from this task indicate that while older adults recall fewer words overall as compared to younger adults, they are just as selective in their memory, remembering a greater proportion of high-value words relative to low-value words. These findings illustrate effective metacognitive monitoring and control in aging as older adults, aware of their limited memory capacity, are able to selectively attend to and remember the high-value words to maximize their score

Further, older adults become more selective after multiple study-test trials when receiving feedback on their memory performance indicating that they are able to incorporate feedback and optimize their encoding strategies (e.g., explicitly allocating more attention towards high-value words) to improve their performance (McGillivray & Castel, 2017). This was especially evident in a variant of the VDR task in which participants were able to self-allocate their study time by selecting which items to study based on their value. In this task, older adults studied high-value words for a greater amount of time relative to younger adults, reducing age-related memory deficits for that high-value information (Castel, Murayama, Friedman, McGillivray, & Link, 2013). Thus, when presented with an excess of information, older adults can accurately assess their memory capacity and adjust their relevant strategies to maximize this capacity, demonstrating successful metacognitive functioning in this VDR task

This finding perhaps reflects a form of selective engagement of cognitive resources (Hess, 2014). As hypothesized, with increasing age, we may require more cognitive effort to engage in particular tasks and our resources may become more quickly depleted when such engagement occurs. As such, older adults may be more selective in the tasks to which they decide to devote cognitive resources, an adaptive response to a reduction in available resources. While Hess (2014) defines this engagement to engage in tasks more broadly, within the specific context of the VDR task, selective engagement may take the form of devoting attentional resources towards high-value information (and inhibiting low-value information) in order to offset reductions in memory capacity experienced with age. This may result in higher selectivity towards high-value items for older adults relative to younger adults, as observed in some VDR studies (Castel, 2008; Castel et al., 2002)

While the previously discussed research suggests that older adults may be able to incorporate knowledge of their memory capacity on a VDR task, little research has investigated the metacognitive monitoring of memory selectivity and how this may change with age. We sought to examine whether predictions related to the amount of information recalled and those related to the ability to selectively recall high-value information may be differentially influenced by the aging process. To our knowledge, no prior work has directly examined older adults' metacognitive judgments of their memory selectivity (i.e., assessing how well participants can engage in selective attention and memory for high-value information) and whether there exist differences in this type of metacognitive functioning and previously studied metacognitive accuracy for memory capacity (i.e., assessing how much information can be remembered). This question represents an important area of research, as it is critical to understand both how accurately older adults can evaluate their ability to remember high-value information and how this accuracy compares to previously established results when examining the metacognitive monitoring of memory capacity. For example, when remembering which medications to take on a given day, it is vital to be able to accurately assess whether you can remember the most important medications relative to less important medications. Even if there are deficits in how older adults evaluate their memory capacity, the ability to evaluate how effectively one can prioritize information in memory is useful and may serve to partially offset those declines

Theoretically, as there exists a dissociation between memory capacity (large age-related declines) and selectivity (minimal age-related declines) in aging (for a review, see Castel, McGillivray, & Friedman, 2012), it may be the case that these two cognitive functions are reliant on different mechanisms – one primarily responsible for the encoding, maintenance, and retrieval of information (a memory mechanism) and another responsible for the prioritization of information based on some characteristics (a selectivity mechanism). Further evidence for the distinction of these two mechanisms is provided by studies investigating VDR in specialized populations. While children with attention-deficit/hyperactivity disorder (ADHD) recalled the same amount of information as age-matched controls, they were significantly less selective, indicating a deficit in the strategic encoding of information based on value, despite no apparent differences in memory capacity (Castel, Lee, Humphreys, & Moore, 2011). Similar work has suggested that patients with Alzheimer’s disease can also selectively remember high-value information despite deficits in recall capacity, although the selectivity may be impaired to some extent, especially for later stage patients (Castel, Balota, & McCabe, 2009). From a neurocognitive perspective, these mechanisms may be reliant on different brain regions, as memory capacity may depend on hippocampal areas typically associated with memory consolidation, while selectivity may depend on prefrontal regions typically responsible for executive functioning including goal maintenance and strategy execution (Cohen, Rissman, Suthana, Castel, & Knowlton, 2014, 2016). It is certainly the case that cognitively healthy older adults do experience impairments in processes like executive functioning and goal maintenance, attributed to early declines in the volume and functioning of prefrontal regions (Glisky, 2007; MacPherson, Phillips, & Della Sala, 2002; Raz et al., 1997; West, 1996). Other work, however, has highlighted the maintained role of reward-based brain regions in old age. Age-invariant activity in the midbrain and ventral striatal regions during value-based tasks may also then contribute to preserved selectivity in a more bottom-up fashion (Cox, Aizenstein, & Fiez, 2008; Rademacher, Salama, Gründer, & Spreckelmeyer, 2014; Samanez-Larkin, Worthy, Mata, McClure, & Knutson, 2014; Spaniol, Bowen, Wegier, & Grady, 2015; Spaniol, Schain, & Bowen, 2014)

Given that memory capacity and selectivity appear to be dissociable to some extent, the ability to monitor and control these processes may also be differentially affected in older age. In the current study, there were two plausible hypotheses regarding the differences between younger and older adults’ varying forms of metacognitive accuracy. Firstly, given that older adults often show deficits in verbal memory capacity, but not in the ability to engage in selective study strategies relative to younger adults (Castel, 2008; Castel et al., 2002; McGillivray & Castel, 2017), older adults may be equally as metacognitively accurate when predicting their ability to be selective, while they may be less accurate when predicting their memory capacity. In this sense, metacognitive functioning may mirror performance on the VDR task, with deficits relative to younger adults in memory capacity predictions, but no differences in memory selectivity predictions (or even older adult superiority in selectivity), suggesting that the underlying mechanisms for these types of predictions may be different (i.e., one that monitors and controls memory capacity and one that monitors and controls memory selectivity) and thus differentially affected by aging. This finding would provide evidence for a dissociation between memory capacity and selectivity such that, while aging may negatively affect the ability to assess memory capacity, it may not affect metacognitive judgments of selectivity. If observed, these results would provide nuance to theories of metacognition and aging, such that the accuracy of global predictions in old age may depend on the type of information being predicted. While metacognitive processes may be impaired when monitoring and controlling memory capacity, such deficits may not exist when assessing one's ability to prioritize information in memory

Alternatively, there may be no differences in metacognitive accuracy between capacity and selectivity for older adults, as these two types of predictions may be reliant on the same underlying mechanism that may be similarly affected by the aging process. We tested these hypotheses in two experiments: Experiment 1 in which younger and older adults predicted how many words they would recall prior to each of four VDR lists and Experiment 2 in which they predicted how many points they would earn on each test

Experiment 1

In Experiment 1, we first sought to investigate how younger and older adults' predictions about their memory capacity in a VDR task may differ. Participants were presented with a series of four unique 20-word lists and were asked to provide a prediction of how many words they would recall at test before each list. They then completed four study-test trials of VDR lists consisting of words paired with point values and were asked to maximize their point score on each list. The inclusion of multiple trials was motivated by prior research that has consistently demonstrated that participants may not optimally execute a value-based study strategy on the first trial, but increase their selectivity towards high-value information with continued task experience and feedback (Castel, 2008; Middlebrooks et al., 2017; Siegel & Castel, 2018a)

We predicted that, consistent with prior research, older adults may recall less information overall than younger adults, but may be comparably selective in the information recalled (Castel, 2008; Castel et al., 2002; McGillivray & Castel, 2017). In terms of metacognitive accuracy, we expected younger adults to be more accurate in their recall predictions than older adults, as younger adults' superior memory capacity performance may also result in more accurate metacognitive knowledge about their memory abilities

Method

Participants—The participants in Experiment 1 were 24 younger adults (M_{age} = 20.08, SD_{age} =1.71, 17 females) and 24 older adults (M_{age} = 77.38, SD_{age} =8.08, 11 females). Younger adults were University of California, Los Angeles (UCLA) undergraduate students who participated for course credit. Older adults were recruited from the local community and compensated $10 per hour, plus parking expenses. Younger adults had completed an average of 13.67 years of education (SD = 1.37), while older adults had completed an average of 16.00 years of education (SD = 2.08). All older adult participants were in self-reported good health and did not report any significant visual impairment

To determine the sensitivity of our analyses with the given sample size, we used the G*Power program. For the later analyses of variance including the relevant parameters (two between-subjects groups and four within-subjects measures) and a power level of 0.8, the resultant effect size was Cohen's f = .17, suggesting that this is the smallest effect that we could have reliably detected with the current sample size. Converting this Cohen's f to eta-squared results in $\eta^2 = .03$ (Cohen, 1988). In both experiments, all significant findings surpassed this value, while all insignificant findings fell below it, suggesting that our sample size provided adequate power to detect significant differences in the following analyses

Materials—The materials utilized in the current study consisted of four lists of 20 words each. The lists contained words ranging in length from four to seven letters (M = 4.99, SD = 0.98) which represented concrete nouns and verbs (e.g., *axle, journal, ride*). On the log-transformed Hyperspace Analogue to Language (HAL) frequency scale (Balota et al., 2007) with lower values indicating lower frequency in the English language and higher values indicating higher frequency, the words ranged from 5.48 to 12.65 and averaged a score of 8.81 (SD = 1.57). In order to avoid specific item effects, for each participant 80 words were randomly drawn from a larger pool of 280 words. This pool was the same as those used in Middlebrooks, Kerr, and Castel (2017). The 80 selected words were then randomly assigned to one of four lists and were randomly paired with a point value from 1 to 10, with two words given each point value. The order of words within each list was also completely randomized for each participant. The result was four lists of 20 words each (two 1-point items per list, two 2-point items, etc.) with a randomized order of point values. As such, word selection, list placement, point value allocation, and point value order were completely randomized for each participant. That is, while one participant may have been presented with the word "axle" on List 1 worth 3-points in the fourth serial position, another participant may have been presented with "axle" on List 3 worth 9-points in the thirteenth serial position. Further, a third participant may have not been presented with the word "axle"

Procedure—Participants were instructed that they would be presented with a series of four 20-word lists. They were then instructed that each word would be paired with a point value ranging from 1 to 10 indicated by a number presented next to each word (and that there would be two words for each point value). The participants' goal was to remember as many of the words in each list as possible while also maximizing their point score (a summation of the points associated with correctly remembered words). They were instructed that after they were presented with a particular list, they would be required to remember only the words from that list (i.e., not previous lists). In Experiment 1, participants were asked to predict how many words they would recall on the upcoming list prior to the presentation of each list with the following question: "Out of 20 words, how many do you think you will remember during the test for this upcoming list?" After participants input their prediction, they were presented with the first randomized list, with each word-point value pair (e.g., axle: 3) being presented in a sequential manner for 3 s each (the total study time for all 20 words was 60 s)

When the study time elapsed after the final word was presented, participants were immediately asked to recall as many words as possible from the previous list. They were instructed that they were not required to input the values associated with the words, just the words themselves. The testing phase was not time-constrained. After participants recorded their responses, they were given feedback on their memory performance. That is, participants were told the number of words (out of 20) that they correctly recalled, but not their total point score. This procedure repeated for the following three lists (for a total of four study-test trials). The experiment concluded when participants received feedback on their performance on the fourth and final list. All materials and procedures used in the current study were approved by the UCLA Institutional Review Board

Results

We first analyzed overall memory accuracy using analyses of variance (ANOVA). Then, in order to examine the effects of item value and task experience on these measures, we used multilevel modeling. Explained in more detail at the below, multilevel modeling is a powerful technique that allowed us to examine the relationship between our variables (i.e., the relationship between item value and recall probability for any given word, and how age group and task experience may have changed this probability). This technique has been used in prior work as a useful analytical approach (Middlebrooks, Murayama, & Castel, 2016; Middlebrooks, Murayama, & Castel, 2017; Siegel & Castel, 2018a, 2018b). However, it does not provide any comparison directly examining mean condition differences (e.g., differences in the overall averages between age groups). In contrast, a mean-based analytic technique (e.g., ANOVA) is unable to detect any direct relationships between item value and recall probability, but is able to determine whether there were differences between age groups on average. As such, the utilization of these analyses in conjunction allowed us to appropriately examine differences in overall recall (using analyses of variance) and differences in selectivity between conditions (using multilevel modeling)

Metacognitive accuracy was assessed using multiple measures. Firstly, to examine group-level differences in accuracy, prediction minus performance scores were calculated for each age group. As such, positive differences indicated overconfidence in group performance, negative differences indicated underconfidence in performance, and difference scores of zero represented accurate predictions of performance. Then, to assess metacognitive accuracy on an individual level, each prediction was correlated with performance on each list for each participant. For this measure, correlations closer to +1 indicated more accurate predictions (i.e., predicting the exact number of words later remembered), while correlations closer to 0 indicated a lack of relationship between predictions and performance. By using both of these measures, we were able to examine both group-level and individual-level differences in prediction accuracy between age groups.

Overall recall and recall predictions—To examine participants' overall memory across the task (depicted in Figure 1), we conducted a 2 (Age group: younger adults, older adults) x 4 (List: 1, 2, 3, 4) repeated-measures analysis of variance (ANOVA) on the number of words recalled. We found a main effect of age group such that younger adults ($M = 6.89$, $SD = 2.03$) recalled more words overall than older adults ($M = 4.66$, $SD = 2.01$), $F(1, 46) = 26.25$, $p < .001$, $\eta^2 = .37$. There was no main effect of list and no interaction between age group and list (ps > .33). This finding indicates that younger adults recalled more information overall than older adults and that both groups of participants recalled a consistent amount of information throughout the task

Next, we examined whether participants' predictions without regard to their actual memory performance (depicted in Figure 2) varied as a function of age group or list by conducting a 2 (Age group: younger adults, older adults) x 4 (List: 1, 2, 3, 4) repeated-measures ANOVA on participants' recall predictions. We found a main effect of list, $F(3, 135) = 12.16$, $p < .001$, $\eta^2 = .21$. Follow-up comparisons with a Bonferroni correction revealed that predictions were higher on List 1 ($M = 8.32$, $SD = 2.91$) as compared to List 2 ($M = 7.23$, $SD = 2.30$), $t(46) = 3.01$, $p = .02$, List 3 ($M = 6.64$, $SD = 1.85$), $t(46) = 4.65$, $p < .001$, and List 4 ($M = 6.28$, $SD = 1.95$), $t(46) = 5.64$, $p < .001$. No other comparisons were significant. The results for the main effect of age group and the interaction between age group and list did not reach standard levels of significance (*p*s > .06). These analyses suggest that both groups of participants predicted that they would recall less information on later lists

Group prediction-performance differences—Of particular interest in the current study was the degree to which participants were accurate in their predictions of their performance (i.e., their metacognitive accuracy). Metacognitive accuracy was measured in two ways: at a group-level examining prediction-performance differences scores and at an individual-level using prediction-performance correlations. We first examined participants' group-level metacognitive accuracy for their memory capacity by calculating recall prediction-performance difference scores (i.e., the number of words that participants predicted that they would recall minus the number of words they actually recalled) depicted in Figure 3. A prediction-performance difference score greater than zero indicates overconfidence in one's memory, while a prediction-performance difference score less than zero indicates underconfidence in one's memory. A difference score of zero indicates perfect accuracy (i.e., participants predicted the same number of words that they later recalled). We conducted a 2 (Age group: younger adults, older adults) x 4 (List: 1, 2, 3, 4) repeated-measures ANOVA on prediction-performance difference scores and found a main effect of age group indicating that younger adults ($M = 0.44$, $SD = 2.81$) were significantly more accurate than older adults ($M = 2.24$, $SD = 2.93$), $F(1, 45) = 11.48$, $p = .001$, $\eta^2 = .20$. Further, we found a main effect of list, $F(3, 135) = 4.66$, $p = .01$, $\eta^2 = .09$. Follow-up comparisons with a Bonferroni correction found that difference scores were significantly higher on List 1 ($M = 2.22$, $SD = 4.08$) than on List 3 ($M = 0.85$, $SD = 2.56$), $t(46) = 2.70$, $p = .05$, and List 4 ($M = 0.54$, $SD = 2.22$), $t(46) = 3.29$, $p = .01$. No other comparisons were significant. There was also no interaction between age group and list ($p = .26$)

To determine whether prediction-performance difference scores were significantly different than zero (indicating perfect accuracy), we conducted one-sample t-tests on the difference scores for List 1–2 and List 3–4 for each age group (lists were grouped in this manner as the prior analysis suggested that Lists 1–2 and Lists 3–4 did not differ significantly in terms of difference scores). Younger adults' prediction-performance difference scores were not significantly different that zero on Lists 1 and 2 ($M = 0.83$, $SD = 2.56$), $t(23) = 1.60$, $p = .12$, or on Lists 3 and 4 ($M = 0.04$, $SD = 1.85$), $t(23) = 0.11$, $p = .91$, indicating highly accurate predictions. On the other hand, older adults' difference scores were significantly greater than zero for both Lists 1 and 2 ($M = 3.13$, $SD = 2.74$), $t(22) = 5.49$, $p < .001$, and Lists 3 and 4 ($M = 1.35$, $SD = 1.94$), $t(22) = 3.33$, $p = .003$. These group-level prediction-performance difference score analyses indicate that while both groups of participants became more accurate in their metamemory with increased task experience, only younger adults were ultimately accurate in their predictions, while older adults remained overconfident throughout the task

Individual prediction-performance correlations—Next, to examine metacognitive accuracy at an individual-level, we computed prediction-performance Pearson's correlations within each age group. That is, each participants' prediction for their memory capacity on List 1 was correlated with their subsequent memory performance on List 1, their prediction for List 2 correlated with their subsequent performance on List 2, and so on for all four lists. Averaged across all four lists, for both younger and older adults, these correlations depicted in Figure 4 were not significant, $r = .15$, $p = .12$, and $r = .11$, $p = .27$, respectively. The magnitude of these correlations was also not significantly different between age groups, $z = 0.28$, $p = .78$

We also calculated correlations for participants' predictions on a list with their previous performance (i.e., the prediction for List 2 correlated with performance on List 1, the prediction for List 3 correlated with performance on List 2, and the prediction for List 4 correlated with performance on List 3). Previous work has shown that predictions may be more highly correlated with performance on the previous list than on subsequent lists, even without explicit feedback given to participants (Hertzog, Dixon, & Hultsch, 1990). In the current task, these correlations may be particularly strong given that explicit feedback was provided after the completion of each trial. Averaged across all four lists, the correlation between predictions and previous recall performance was significantly positive for both younger, $r = .48$, $p < .001$, and older adults, $r = .39$, $p < .001$. Comparing between the magnitude of these coefficients suggests there was no significant difference, z = 0.26, $p = .52$. So, when examining this measure of individual metacognitive accuracy, neither younger nor older adults' predictions were significantly associated with their subsequent performance across the task, but were both significantly positively correlated with performance on the previous list. Thus, these results suggest that both younger and older adults were somewhat relying on previous task performance to the same extent to make predictions about recall performance on upcoming lists

Memory selectivity—Participants' overall recall with regards to item value is illustrated in Figure 5. In order to compare selectivity between groups and across lists, we used multilevel modeling to model the number of words recalled as a function of item value. Multilevel modeling has been used in previous studies investigating memory selectivity (Castel et al., 2013; Middlebrooks & Castel, 2017; Middlebrooks et al., 2016, 2017; Raudenbush & Bryk, 2002; Siegel & Castel, 2018a, 2018b). The post-hoc binning of items into low, medium, and high value groups (as would be needed in an ANOVA) may not accurately reflect participants' valuations of to-be-learned stimuli (e.g., Participant 1 may consider items with values 6–10 to be of "high" value, while Participant 2 may only consider items with values 8–10 as such). In contrast, multilevel modeling treats item value as a continuous variable, allowing for a more precise investigation of the relationship between the number of words recalled and item value. Further, by first clustering data within each participant and then examining possible condition differences, multilevel modeling accounts for both within- and between-subject differences in strategy use, the latter of which would not be evident when conducting standard analyses of variance. Thus, multilevel modeling allows for a more fine-grained analysis of participants' value-based strategies

In a two-level model, recall probability (using a Bernoulli distribution, 0 = not recalled, 1 = recalled; level 1 = items; level 2 = participants) was modeled as a function of item value, list, and the interaction between those two variables. Item value and list were entered into the model as group-mean centered variables (with item value anchored at the mean value of 5.5 and list anchored at the mean value of 2.5). The age groups (0 = older adults, 1 = younger adults) were included as level-2 predictors. Regression coefficients (β) obtained from multilevel model can be interpreted via their exponential (Raudenbush & Bryk, 2002) – that is, the Exp(β) represents the effect of the independent variable on the odds ratio of recall probability (the probability of successful recall divided by the unsuccessful recall probability). An Exp(β) value greater than one indicates a positive effect of a predictor, while an Exp(β) value less than one indicates a negative effect of a predictor

Firstly, there was a significant effect of value on recall probability for older adults, β_{10} = 0.23, $p < .001$. This indicates that for each increase in item value, older adults were $e^{0.23}$ = 1.26 times more likely to correctly remember that item. Further, older adults were $e^{0.23*10}$ = 10.17 times more likely to successfully to remember a 10-point item, as compared to a 1-point item. Thus, as item value increased, older adults were more likely to accurately recall the items. However, this effect was significantly different for younger adults, $\beta_{11} = -0.11$, $p = .04$. To calculate the simple slope for younger adults, the β_{10} and β_{11} coefficients were added ($\beta_{YAs} = 0.11$). To determine the significance of these slopes, the model was adjusted to treat younger adults as the comparison group (0 = younger adults, 1 = older adults). This method was used through the remainder of the analyses in the current study to calculate the significance of simple slopes. The adjusted analysis revealed that item value was in fact a significant predictor of recall probability for younger adults, $\beta_{YAs} = 0.11$, $p = .004$. That is, for each increase in item value, younger adults were $e^{0.11}$ = 1.12 times more likely to recall an item and $e^{0.11*10}$ = 3.10 times more likely to recall a 10-point relative to a 1-point item. Taken together, these results suggest that while both younger and older adults were selective towards high-value information across lists, older adults were significantly more selective than younger adults. That is, the positive relationship between item value and recall probability was stronger for older relative to younger adults, indicating a higher level of selectivity in their recall

Secondly, list was not a significant predictor of recall probability for older adults, $\beta_{20} = -0.02$, $p = .74$, which was not significantly different for younger adults, $\beta_{21} = -0.07$, $p = .29$, suggesting that both groups of participants recalled the same amount of information across lists (regardless of item value). Finally, there was a significant positive interaction between item value and list for older adults, $\beta_{30} = 0.05$, $p = .02$, which was not significantly different for younger adults, $\beta_{31} = -0.02$, $p = .59$. This suggests that the relationship between item value and recall probability increased across lists. That is, both younger and older adults became more selective towards high-value information with increased task experience

Discussion

The results from Experiment 1 suggest that while both groups were selective towards high-value words, older adults were more selective than younger adults in the information that they recalled. However, as indicated by group-level prediction-performance difference, older adults' ability to predict their memory capacity may be less accurate than their younger adult counterparts. Although older adults became more metacognitively accurate with increased task experience (indicated by decreasing difference scores across lists), results indicated that they were still overconfident in their memory capacity after multiple study-test trials (indicated by difference scores still significantly greater than zero on Lists 3–4), consistent with prior research examining global predictions of memory capacity (Connor et al., 1997; Hertzog et al., 1994). These results are consistent with previous work which found that older adults' metacognitive performance on a VDR task improved with experience due to the amount of information "bet on" by older adults decreasing paired with a consistent level of recall (McGillivray & Castel, 2011). Younger adults, on the other hand, were perfectly accurate from the beginning of the task, and remained so throughout. Importantly, for individual-level measures of metacognitive accuracy, neither younger nor older adults' predictions of their memory capacity were correlated with their subsequent performance on the next list, suggesting that these groups may be equally inaccurate in predicting their own performance when using this measure

Experiment 2

As Experiment 1 revealed age-related differences in the metacognitive accuracy of memory capacity on a VDR task, we sought to determine whether these differences would be present when participants were asked to assess how well they could selectively remember information. In Experiment 2, new groups of younger and older adults were asked to predict how many points they would earn out of 110 possible prior to each VDR list. In addition to similar recall and selectivity effects found in Experiment 1, we expected that, in contrast to recall predictions, there would be no difference in the accuracy of points predictions between younger and older adults. As older adults may be comparably selective in their memory relative to younger adults on these VDR tasks, this may also result in equivalent metacognitive accuracy for this type of information

Method

Participants—The participants in Experiment 2 were 24 younger adults (M_{age} = 20.58, SD_{age} =2.48, 20 females) and 24 older adults (M_{age} = 75.75, SD_{age} = 6.92, 9 females). Younger adults were UCLA undergraduate students who participated for course credit. Older adults were recruited from the local community and compensated $10 per hour, plus parking expenses. Younger adults had completed an average of 13.92 years of education (*SD* = 1.22), while older adults had completed an average of 16.08 years of education (*SD* = 2.27). All older adult participants were in self-reported good health and did not report any significant visual impairment. Participants in Experiment 1 were excluded from participation in Experiment 2

Materials and procedure—The materials were identical to those utilized in Experiment 1 (i.e., the same pool of 280 words was used to form the four unique 20-word lists for each participant). The procedure was also identical to Experiment 1 with one key exception: prior to each of the four lists, participants were asked to make predictions about their point total on the upcoming list (as opposed to the number of words they would recall in Experiment 1). That is, participants were asked "Out of 110 points, how many do you think you will earn on this upcoming list?" Participants were then presented with the 20-word list and asked to recall as many words as possible. They were then given feedback on their point score, as they were informed of the number of points (out of 110) that they earned on the current list, but not on the number of words they recalled. Participants then repeated this procedure for the remaining three lists (for a total of four study-test trials)

Results

Overall recall and points predictions—Similar to Experiment 1, overall memory across the task (shown in Figure 6) was examined using a 2 (Age group: younger adults, older adults) x 4 (List: 1, 2, 3, 4) repeated-measures analysis of variance (ANOVA) on the number of words recalled. We found a main effect of age group such that younger adults ($M = 8.26$, $SD = 2.62$) recalled more words overall than older adults ($M = 4.84$, $SD = 1.71$), $F(1, 46) = 28.65$, $p < .001$, $\eta^2 = .38$. There was no main effect of list and no interaction between age group and list (ps > .22). As in Experiment 1, this indicates that younger adults recalled more words overall than older adults and that both groups participants recalled a consistent amount of words throughout the task

Next, to examine whether participants' points predictions (depicted in Figure 7) varied as a function of age group or list we conducted a 2 (Age group: younger adults, older adults) x 4 (List: 1, 2, 3, 4) repeated-measures ANOVA on participants' points predictions (ranging from 0 to 110 per list). Firstly, there was a main effect of age group such that younger adults ($M = 52.58$, $SD = 18.50$) predicted that they would earn more points than older adults ($M = 33.25$, $SD = 15.05$), $F(1, 46) = 31.82$, $p < .001$, $\eta^2 = .41$. We also found a main effect of list, $F(3, 138) = 10.83$, $p < .001$, $\eta^2 = .19$. Follow-up comparisons with a Bonferroni correction revealed that predictions were higher on List 1 ($M = 51.96$, $SD = 22.39$) as compared to List 2 ($M = 40.46$, $SD = 19.27$), $t(46) = 4.42$, $p < .001$, List 3 ($M = 40.15$, $SD = 16.65$), $t(46) = 4.54$, $p < .001$, and List 4 ($M = 39.10$, $SD = 16.41$), $t(46) = 4.94$, $p < .001$. No other follow-up comparisons were significant. There was no interaction between age group and list, $F(3, 138) = 0.83$, $p = .48$, $\eta^2 = .01$. These analyses suggest that younger adults predicted they would earn more points than older adults overall and that both groups of participants predicted that they would recall less information on later lists

Group prediction-performance differences—We first examined participants' metacognitive accuracy for their ability to be selective by calculating prediction-performance difference (i.e., the number of points that participants predicted that they would earn minus the number of points they actually earned) at a group-level which are portrayed in Figure 8. We conducted a 2 (Age group: younger adults, older adults) x 4 (List: 1, 2, 3, 4) repeated-measures ANOVA on prediction-performance difference scores and found no main effect of age group indicating that younger adults ($M = 3.78$, $SD = 11.70$) were equally as accurate in their predictions as older adults (M = 1.38, SD = 7.45) across lists, $F(1, 46) = 0.72$, $p = .40$, $\eta^2 = .02$. We found a main effect of list, $F(3, 138) = 8.37$, $p < .001$, $\eta^2 = .15$. Follow-up comparisons with a Bonferroni correction found that prediction-performance difference scores were significantly higher on List 1 (M = 14.19, SD = 27.16) than on List 2 (M = - 0.04, SD = 14.11), $t(46) = 3.70$, $p = .002$, List 3 (M = −0.63, SD = 15.53), $t(46) = 3.89$, $p = .001$, and List 4 (M = −3.21, SD = 15.90), $t(46) = 4.53$, $p < .001$. No other comparisons were significant. There was no interaction between age group and list, $F(3, 138) = 0.33$, $p = .80$, $\eta^2 = .01$. Contrary to Experiment 1, these results suggest that younger and older adults were equally as accurate in their predictions of their point scores

Next, we sought to determine whether prediction-performance difference scores were significantly different than zero (with zero indicating perfect accuracy). Therefore, we conducted one-sample t-tests on the difference scores for List 1 and List 2–4 collapsed across age groups. Lists were grouped in this manner as the prior analysis suggested that there was no difference in difference scores between younger and older adults, difference scores on List 1 were significantly higher than the other lists, and no difference existed between prediction-performance difference scores on Lists 2–4. These analyses revealed that difference scores were significantly greater than zero on List 1 (M = 14.19, SD = 27.16), $t(47) = 3.62$, $p < .001$. However, difference scores were not significantly different than zero on Lists 2–4 (M = −1.29, SD = 9.81), $t(47) = 0.91$, $p = .37$. These analyses indicate that both groups of participants became accurate in their point predictions after List 1. This stands in contrast to Experiment 1 where only younger adults were predicting their recall accurately, while older adults were overconfident in their memory performance throughout the task

Individual prediction-performance correlations—We again computed prediction-performance correlations as an individual-level of metacognitive accuracy for both previous and subsequent lists. Pearson's correlations between points predictions and points performance on the subsequent list (e.g., List 1 predictions with List 1 performance) for each age group are depicted in Figure 9. Similar to Experiment 1, there was no significant relationship between these measures for younger adults, $r = .15$, $p = .15$. However, there was a significant positive correlation between points predictions and performance for older adults, $r = .49$, $p < .001$, suggesting that older adult individuals who predicted they would earn more points tended to do so during the task. There was a significant difference in the magnitude of these coefficients between younger and older adults, $z = 2.62$, $p = .01$. Thus, when utilizing this measure of individual metacognitive accuracy, younger adults' predictions were not significantly associated with their performance across the task, while older adults exhibited greater individual-level accuracy

In terms of predictions and performance on the previous list (e.g., List 2 predictions with List 1 performance), there was a significantly positive relationship for both younger, $r = .84$, $p < .001$, and older adults, $r = .77$, $p < .001$, across all four lists. There was no difference in the magnitude of these correlation coefficients, $z = 1.18$, $p = .24$. These results suggest that both younger and older adults' points predictions were heavily reliant on their previous point score. However, only older adults' predictions of points were significantly associated with their subsequent performance across lists

Memory selectivity—We applied the same model described in Experiment 1 to conduct a multilevel model analysis to examine participants' memory selectivity across lists (depicted in Figure 10). That is, in a two-level model, recall probability (using a Bernoulli distribution, 0 = not recalled, 1 = recalled; level 1 = items; level 2 = participants) was modeled as a function of item value, list, and the interaction between those two variables. Firstly, we found that value was a significant positive predictor of recall probability for older adults, $\beta_{10} = 0.22$, $p < .001$. This indicates that for each increase in item value, an item was $e^{0.22} = 1.24$ times more likely to be recalled and that a 10-point item was $e^{0.22*10} = 8.76$ times more likely to be recalled than a 1-point item. This analysis also revealed a marginally significant difference in the effects of item value on recall probability for younger adults, $\beta_{11} = -0.12$, $p = .05$. Conducting the analysis using younger adults as the comparison group revealed that item value was also a positive predictor of recall probability for that group, $\beta_{YAs} = 0.10$, $p < .001$. Similar to Experiment 1, these results suggest that while both groups recalled more high-value than low-value information across lists, older adults were more selective than younger adults in the information they recalled

Secondly, list was not a significant predictor of recall probability for older adults, $\beta_{20} = 0.03$, $p = .55$, which was not significantly different for younger adults, $\beta_{21} = -0.01$, $p = .94$. Finally, there was no significant interaction between item value and list for older adults, $\beta_{30} = .02$, $p = .39$, which was not significantly different for younger adults, $\beta_{31} = -0.01$, $p = .85$. This result differs from Experiment 1, where the relationship between item value and recall probability became more positive across lists.

Discussion

Recall and selectivity results from Experiment 2 were largely consistent with Experiment 1 in that younger adults recalled more information than older adults overall, and selectivity was higher for older adults. However, in Experiment 2 there was no significant difference between younger and older adults in terms of their prediction-performance difference scores, with both groups displaying similar magnitudes of overconfidence on the first study-test trial and accurate predictions thereafter. These results suggest that the age-related differences present when predicting memory were not present when predictions were made about the ability to selectively study information. Further, when examining individual-level metacognitive accuracy via prediction-performance correlations, older adults exhibited superior accuracy than younger adults. This was indicated by a positive correlation suggesting that those older adult individuals who predicted they would earn more points actually did so, while younger adults' predictions were not significantly associated with their actual points performance

General Discussion

While episodic memory capacity tends to decline as we age (Park et al., 2002), some metamemorial processes tend to remain relatively intact (Hertzog & Dunlosky, 2011). As demonstrated by performance on value-directed remembering (VDR) tasks, when presented with an excess of information, older adults are able to accurately assess their memory ability and adjust their study strategies to maximize performance in order to selectively remember high-value information (Ariel et al., 2015; Castel, 2008; Castel et al., 2002; Hayes et al., 2012; McGillivray & Castel, 2017). The current experiments examined whether younger and older adults' metacognitive judgments may differ in accuracy depending on the type of information assessed. Theoretically, a potential dissociation between memory capacity and selectivity is useful, especially if aging may impair the ability to accurately assess memory capacity, but not metacognitive judgments of selectivity. Specifically, we were interested in whether metamemorial judgments on a VDR task would match age-related memory findings – that is, whether there would be deficits among older adults when evaluating memory capacity, but no difference in performance when evaluating memory selectivity

In Experiment 1, younger and older adults assessed their memory capacity by predicting how many words they would remember before each of four VDR study-test trials. Younger adults recalled more words overall, while older adults were significantly more selective in the information recalled. Most importantly, younger adults were accurately predicting the number of words they would later recall from the beginning of the task. Older adults, on the other hand, were overconfident in their predictions throughout the entire task, despite becoming more accurate with task experience. Comparisons of individual-level accuracy revealed no differences in between younger and older adults. In Experiment 2, participants assessed their memory selectivity by predicting how many points they would earn before completing the four VDR study-test trials. Memory and selectivity results were consistent with Experiment 1. Crucially however, there was no significant difference in the accuracy of points predictions between younger and older adults. Both groups of participants were overconfident on the first trial, but were accurately predicting the number of points they would later earn on subsequent trials. Further, individual-level prediction-performance correlations indicated older adults were more accurate in predicting their points performance on the upcoming trial

The current experiments replicate previous findings and add further evidence that, despite deficits in memory capacity relative to younger adults, older adults are able to engage in selective study strategies to maximize their performance (Ariel et al., 2015; Castel, 2008; Castel et al., 2002; Hayes et al., 2012; McGillivray & Castel, 2017; Spaniol et al., 2014). In both experiments, younger adults recalled more words than older adults, but selectivity was higher for older adults relative to younger adults (significantly in Experiment 1 and marginally in Experiment 2). These findings suggest that older adults use information about their memory capacity in this VDR task to adjust their goal-relevant strategies by allocating attention towards high-value and away from low-value information, consistent with the idea that older adults are more selective in engaging resources in order to compensate for a reduction in those resources (Hess, 2014). As such, older adults' performance in these experiments demonstrates effective metacognitive control. It is however important to note that, due to the feedback provided to participants, the monitoring aspect of metacognition was relatively equivalent between age groups, as participants' own monitoring processes were supplemented by the explicit feedback on task performance in both studies (in terms of words recalled in Experiment 1 and points earned in Experiment 2).

Perhaps the most novel findings of the current experiments are overall group-level prediction-performance differences and individual-level prediction-performance correlations observed between Experiment 1 and 2. For group-level differences, older adults in Experiment 1 demonstrated an overconfidence in their memory capacity throughout the task, consistent with prior work utilizing global judgments of performance (Bruce et al., 1982; Connor et al., 1997; Hertzog et al., 1994). Younger adults, on the other hand, were fairly accurate at predicting their own memory capacity. In terms of individual-level correlations, neither group exhibited a significant relationship between their predictions and performance on the subsequent list. However, predictions were associated with performance on the previous list for both age groups, suggesting that they were using their previous recall scores to guide their predictions. In Experiment 2, prediction-performance difference scores did not differ between younger and older adults, suggesting at least equivalent metacognitive accuracy when assessing the ability to selectively remember high-value information in this VDR task. On an individual level, older adults' points predictions were positively correlated with their performance on the subsequent list, while younger adults exhibited no such relationship. Further, both age groups' predictions during Experiment 2 were very highly associated with previous list performance, suggesting that they were relying on the provided feedback to make predictions about upcoming lists, consistent with prior research (Hertzog et al., 1990).

Previous work has shown that participants' predictions may be more highly associated with their previous performance than with subsequent task performance, even though predicting subsequent task performance is the goal of the prediction (Hertzog et al., 1990). Results from the current study replicate this finding and suggest that participants may even more heavily rely on previous task performance to make predictions when explicit feedback is provided. Although not statistically examined due to methodological concerns comparing between experiments, these positive correlations were particularly high in Experiment 2 when predicting how many points they would earn (r_{YA} = .84, r_{OA} = .77) relative to Experiment 1 when predicting how many words they would remember (r_{YA} = .48, r_{OA} = .39). Given that the monitoring of points earned may be more difficult and less intuitive than the monitoring of words remembered, it is likely that participants relied more heavily on previous task performance to make predictions about how many points they would earn on the subsequent list. It is important to note that despite this reliance on previous performance to predict upcoming performance, prior work examining the memory-for-past-test (MPT) heuristic has shown that participants may also rely on factors other than previous test performance to make predictions (Ariel & Dunlosky, 2011; Hertzog, Hines, & Touron, 2013; Hines, Hertzog, & Touron, 2015; Tauber & Rhodes, 2012). Factors such as the new learning, forgetting, and subjective confidence may also have effects on performance monitoring suggesting that multiple complex cues influence monitoring (Ariel & Dunlosky, 2011; Hertzog et al., 2013). Further, as replicated by the results in the current study, younger and older adults may rely on the MPT heuristic to the same extent to guide predictions (Hines et al., 2015; Tauber & Rhodes, 2012). The current study also suggests that this reliance on the MPT heuristic may be exacerbated when judgments are made about less intuitive types of information, such as the number of points one will earn.

More interestingly, there were significant differences between younger and older adults in terms of their predictions and their subsequent performance. While participants were clearly relying on previous task performance to make predictions about upcoming performance, their goal as specified in the instructions was to accurately predict their *subsequent* memory performance. As such, the analyses examining correlations between predictions and subsequent list performance are more representative of task goals than correlations between predictions and previous list performance. In Experiment 1, for prediction-subsequent list performance correlations, there were no age differences when predicting words. However, in Experiment 2, older adults' points predictions were significantly positively associated with upcoming points performance, while younger adults' predictions were unrelated to their points performance. These results suggest that older adults were *more* accurate when predicting how many points they would earn on the next list relative to younger adults. In contrast, younger and older adults were equally as accurate (or inaccurate, rather) in predicting their upcoming recall performance.

Taken together, these findings represent a dissociation – that is, age-related deficits were present when predicting memory capacity, but absent when predicting memory selectivity (or even more accurate performance on some measures for older adults). This mirrors actual performance on the VDR task in the current study and previous work, such that older adults exhibit superior or equivalent selectivity despite decrements in the amount of information recalled (Castel et al., 2002, 2012; Castel, 2008; McGillivray & Castel, 2017). This finding adds nuance to theories of metacognition and aging by suggesting that the accuracy of predictions in old age may depend on the type of information being assessed. Accordingly, these findings may provide potential evidence for separate underlying metacognitive mechanisms – one that monitors and controls memory capacity (which appears to become impaired with age) and one that monitors and control memory selectivity (which appears to be relatively unaffected by aging).

It is important to note that while global predictions of performance in the task likely represent the monitoring of item-level performance to some extent, they may also be based on other factors including individual goal-setting and self-efficacy, as well as implicit theories about possible ranges of performance that participants may bring to the task. With this limitation in mind, these global predictions are not unbiased measures of metacognitive monitoring, but are likely influenced by a multitude of other factors. As such, future research is needed to further examine how metacognitive functioning may differ depending on the type of information being assessed. For example, while global predictions were made in the current experiments, as discussed, it may be informative to explore how metacognitive accuracy may differ when judgments are made on an item-by-item basis (cf. Halamish et al., 2011) in order to make more direct claims about item-level metacognitive monitoring. In real-world situations, people may evaluate how well they know each piece of information on an individual item basis as it is encountered (e.g., how likely am I to remember to take this medication later?), as opposed to making a pre-task global judgment about how likely the information is to remembered (e.g., how likely am I to remember all of the things I must get done today?).

For example, McGillivray and Castel (2011) found that, using a VDR task, when asked to bet on words they would later recall, older adults were overconfident in their performance, betting on more words than they could remember. However, both younger and older adults were selective in their betting (i.e., betting more on high-value words) and their recall (i.e., recalling those high-value words), consistent with results from the current study that demonstrate overconfidence in predicting memory capacity, but competency in predicting memory selectivity. As such, these results suggest that older adults may be metacognitively accurate when betting on an individual item-by-item basis in this task. Future studies should consider directly comparing individual item-by-item judgments with global predictions and how these different types of metacognitive processes may vary with age and the value of information. Further, it would be useful for future research to explore how these hypothesized separate metacognitive mechanisms may compare within-participant. In the current experiments, two separate groups of participants were used to compare metacognitive accuracy of global predictions. For example, if separate mechanisms do exist, then within the same older adult participant, we would expect accurate monitoring and control of selectivity, but not global predictions of memory capacity. These findings would provide further direct evidence for the dissociation of these mechanisms and thus merits future investigation.

The current study investigated how metacognitive judgments of memory capacity and the ability to selectivity remember information may change as we age. In both experiments, younger adults recalled more information overall and older adults were more selective in the information they remembered. However, the accuracy of metacognitive judgments (measured via pre-study predictions) varied depending on the type of information being assessed. While older adults were less accurate in predicting their memory capacity displaying overconfidence, they were equally as accurate (or even more accurate when examining individual-level correlations) when predicting the amount of points they would earn. These findings suggest that older adults were able to effectively monitor and control their ability to selectively remember high-value information, despite recalling less information than younger adults. In sum, these results provide further evidence of a dissociation between memory capacity and selectivity and demonstrate that, while age may impair the ability to assess memory capacity, it may not affect metacognitive judgments of selectivity.

Acknowledgements

The authors would like to thank Adam Blake, Brandon Carone, Mary Hargis, Tyson Kerr, Catherine Middlebrooks, and the rest of the members of the Memory and Lifespan Cognition Laboratory at UCLA for their guidance and support throughout this project.

This research was supported in part by the National Institutes of Health (National Institute on Aging; Award Number R01 AG044335 to Alan Castel). Portions of this work were presented at the 2018 Cognitive Aging Conference in Atlanta, GA and the UCLA Research Conference on Aging in Los Angeles, CA.

References

Ariel R, & Dunlosky J (2011). The sensitivity of judgments-of-learning resolution to past test performance, new learning, and forgetting. Memory & Cognition, 39, 171–184. [PubMed: 21264621]

Ariel R, Price J, & Hertzog C (2015). Age-related associative memory deficits in value-based remembering: The contribution of agenda-based regulation and strategy use. Psychology and Aging, 30, 795–808. [PubMed: 26523692]

Balota DA, Yap MJ, Cortese MJ, Hutchison KA, Kessler B, Loftis B, ... Treiman R (2007). The English lexicon project. Behavior Research Methods, 39, 445–459. [PubMed: 17958156]

Bruce PR, Coyne AC, & Botwinick J (1982). Adult age differences in metamemory. Journal of Gerontology, 37, 354–357. [PubMed: 7069161]

Castel AD (2008). The adaptive and strategic use of memory by older adults: Evaluative processing and value-directed remembering In Benjamin AS & Ross BH (Eds.), The psychology of learning and motivation (Vol. 48, pp. 225–270). London: Academic Press.

Castel AD, Balota DA, & McCabe DP (2009). Memory efficiency and the strategic control of attention at encoding: Impairments of value-directed remembering in Alzheimer's disease. Neuropsychology, 23, 297–306. [PubMed: 19413444]

Castel AD, Benjamin AS, Craik FIM, & Watkins MJ (2002). The effects of aging on selectivity and control in short-term recall. Memory & Cognition, 30, 1078–1085. [PubMed: 12507372]

Castel AD, Lee SS, Humphreys KL, & Moore AN (2011). Memory capacity, selective control, and value-directed remembering in children with and without attention-deficit/hyperactivity disorder. Neuropsychology, 25, 15–24. [PubMed: 20873928]

Castel AD, McGillivray S, & Friedman MC (2012). Metamemory and memory efficiency in older adults: Learning about the benefits of priority processing and value-directed remembering In Naveh-Benjamin M & Ohta N (Eds.) Memory and aging: Current issues and future directions (pp. 245–270). New York: Psychology Press.

Castel AD, Middlebrooks CD, & McGillivray S (2016). Monitoring memory in old age: Impaired, spared, and aware In Dunlosky J & Tauber S (Eds.), The Oxford handbook of metamemory (pp. 519–536). Oxford: Oxford University Press.

Castel AD, Murayama K, Friedman MC, McGillivray S, & Link I (2013). Selecting valuable information to remember: Age-related differences and similarities in self-regulated learning. Psychology and Aging, 28, 232–242. [PubMed: 23276210]

Cohen J (1988). Statistical power analysis for the behavioral sciences, 2nd ed. Hillsdale, NJ: Lawrence Earlbaum.nd

Cohen MS, Rissman J, Suthana NA, Castel AD, & Knowlton BJ (2014). Value-based modulation of memory encoding involves strategic engagement of fronto-temporal semantic processing regions. Cognitive, Affective, & Behavioral Neuroscience, 14, 578–592.

Cohen MS, Rissman J, Suthana NA, Castel AD, & Knowlton BJ (2016). Effects of aging on value-directed modulation of semantic network activity during verbal learning. NeuroImage, 125, 1046–1062. [PubMed: 26244278]

Connor LT, Dunlosky J, & Hertzog C (1997). Age-related differences in absolute but not relative metamemory accuracy. Psychology and Aging, 12, 50–71. [PubMed: 9100268]

Cox KM, Aizenstein HJ, & Fiez JA (2008). Striatal outcome processing in healthy aging. Cognitive, Affective, & Behavioral Neuroscience, 8, 304–317.

Glisky EL (2007). Changes in cognitive function in human aging In Riddle DR (Ed.), Brain aging: Models, methods, and mechanisms (pp. 3–20). Boca Raton, FL: CRC Press.

Halamish V, McGillivray S, & Castel AD (2011). Monitoring one's own forgetting in younger and older adults. Psychology and Aging, 26, 631–635. [PubMed: 21463057]

Hayes MG, Kelly AJ, & Smith AD (2012). Working memory and the strategic control of attention in older and younger adults. Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 68, 176–183.

Hertzog C, & Dixon RA (1994). Metacognitive development in adulthood and old age In Metcalfe J & Shimamura A (Eds.), Metacognition: Knowing about knowing (pp. 227–251). Cambridge, MA: Bradford.

Hertzog C, Dixon RA, & Hultsch DF (1990). Relationships between metamemory, memory predictions, and memory task performance in adults. Psychology and Aging, 5, 215–227. [PubMed: 2378687]

Hertzog C, & Dunlosky J (2011). Metacognition in later adulthood: Spared monitoring can benefit older adults' self-regulation. Current Directions in Psychological Science, 20, 167–173. [PubMed: 24478539]

Hertzog C, Hines JC, & Touron DR (2013). Judgments of learning are influenced by multiple cues in addition to memory for past test accuracy. Archives of Scientific Psychology, 1, 23–32. [PubMed: 25914865]

Author Manuscript
Author Manuscript
Author Manuscript

Hertzog C, Saylor LL, Fleece AM, & Dixon RA (1994). Metamemory and aging: Relations between predicted, actual and perceived memory task performance. Aging, Neuropsychology, and Cognition, 1, 203–237.

Hertzog C, Sinclair SM, & Dunlosky J (2010). Age differences in the monitoring of learning: Cross-sectional evidence of spared resolution across the adult life span. Developmental Psychology, 46, 939–948. [PubMed: 20604613]

Hess TM (2005). Memory and aging in context. Psychological Bulletin, 131, 383–406. [PubMed: 15869334]

Hess TM (2014). Selective engagement of cognitive resources: Motivational influences on older adults' cognitive functioning. Perspectives on Psychological Science, 9, 388–407. [PubMed: 26173272]

Hines JC, Hertzog C, & Touron DR (2015). Younger and older adults weight multiple cues in a similar manner to generate judgments of learning. Aging, Neuropsychology, and Cognition, 22, 693–711.

Hines JC, Touron DR, & Hertzog C (2009). Metacognitive influences on study time allocation in an associative recognition task: An analysis of adult age differences. Psychology and Aging, 24, 462–475. [PubMed: 19485662]

Kavé G, & Halamish V (2015). Doubly blessed: Older adults know more vocabulary and know better what they know. Psychology and Aging, 30, 68–73. [PubMed: 25602490]

MacPherson SE, Phillips LH, & Della Sala S (2002). Age, executive function, and social decision making: A dorsolateral prefrontal theory of cognitive aging. Psychology and Aging, 17, 598–609. [PubMed: 12507357]

McGillivray S, & Castel AD (2011). Betting on memory leads to metacognitive improvement by younger and older adults. Psychology and Aging, 26, 137–142. [PubMed: 21417541]

McGillivray S, & Castel AD (2017). Older and younger adults' strategic control of metacognitive monitoring: The role of consequences, task experience, and prior knowledge. Experimental Aging Research, 43, 233–256. [PubMed: 28358293]

Middlebrooks CD, & Castel AD (2017). Self-regulated learning of important information under sequential and simultaneous encoding conditions. Journal of Experimental Psychology: Learning, Memory, and Cognition, 44, 779–792.

Middlebrooks CD, Kerr T, & Castel AD (2017). Selectively distracted: Divided attention and memory for important information. Psychological Science, 28, 1103–1115. [PubMed: 28604267]

Middlebrooks CD, Murayama K, & Castel AD (2016). The value in rushing: Memory and selectivity when short on time. Acta Psychologica, 170, 1–9. [PubMed: 27305652]

Middlebrooks CD, Murayama K, & Castel AD (2017). Test expectancy and memory for important information. Journal of Experimental Psychology: Learning, Memory, and Cognition, 43, 972–985.

Park DC, Lautenschlager G, Hedden T, Davidson NS, Smith AD, & Smith PK (2002). Models of visuospatial and verbal memory across the adult life span. Psychology and Aging, 17, 299–320. [PubMed: 12061414]

Rademacher L, Salama A, Gründer G, & Spreckelmeyer KN (2014). Differential patterns of nucleus accumbens activation during anticipation of monetary and social reward in young and older adults. Social Cognitive and Affective Neuroscience, 9, 825–831. [PubMed: 23547243]

Raudenbush SW, & Bryk AS (2002). Hierarchical linear models: Applications and data analysis methods (2nd ed.). Newbury Park, CA: Sage.nd

Raz N, Gunning FM, Head D, Dupuis JH, McQuain J, Briggs SD, ... & Acker JD (1997). Selective aging of the human cerebral cortex observed in vivo: Differential vulnerability of the prefrontal gray matter. Cerebral Cortex, 7, 268–282. [PubMed: 9143446]

Samanez-Larkin GR, Worthy DA, Mata R, McClure SM, & Knutson B (2014). Adult age differences in frontostriatal representation of prediction error but not reward outcome. Cognitive, Affective, & Behavioral Neuroscience, 14, 672–682.

Siegel ALM, & Castel AD (2018a). Memory for important item-location associations in younger and older adults. Psychology and Aging, 33, 30–45. [PubMed: 29494176]

Siegel ALM, & Castel AD (2018b). The role of attention in remembering important item-location associations. Memory & Cognition, 46, 1248–1262. [PubMed: 29926393]

Soderstrom NC, McCabe DP, & Rhodes MG (2012). Older adults predict more recollective experiences than younger adults. Psychology and Aging, 27, 1082–1088. [PubMed: 22686405]

Spaniol J, Bowen HJ, Wegier P, & Grady C (2015). Neural responses to monetary incentives in younger and older adults. Brain Research, 1612, 70–82. [PubMed: 25305570]

Spaniol J, Schain C, & Bowen HJ (2014). Reward-enhanced memory in younger and older adults. Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 69, 730–740. [PubMed: 23690000]

Tauber SK, & Rhodes MG (2012). Multiple bases for younger and older adults' judgments of learning in multitrial learning. Psychology and Aging, 26, 474–483.

West RL (1996). An application of prefrontal cortex theory to cognitive aging. Psychological Bulletin, 120, 272–292. [PubMed: 8831298]

Zacks RT, & Hasher L (2006). Aging and long-term memory: Deficits are not inevitable In Bialystok E & Craik FIM (Eds.), Lifespan cognition: Mechanisms of change (pp. 162–177). Oxford: Oxford University Press.

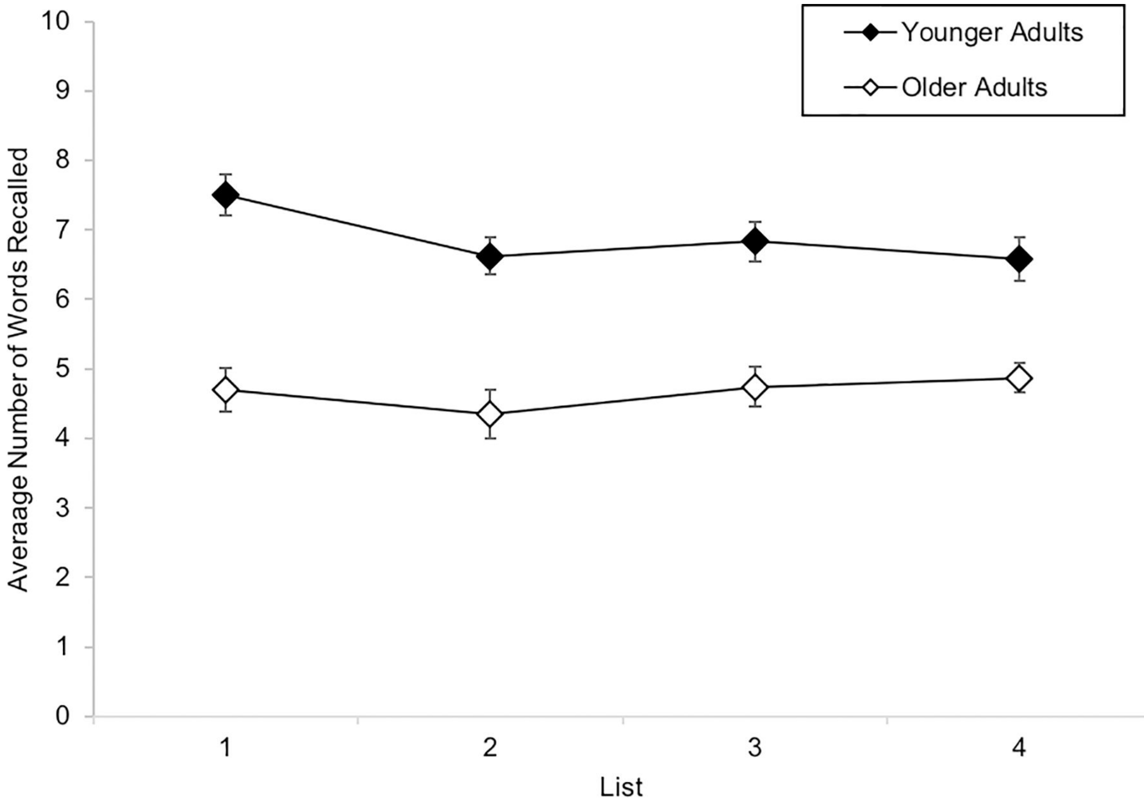

Figure 1.
Average number of words recalled for both age groups across lists (out of 20 possible words) in Experiment 1. Error bars represent ± 1 standard error of the mean.

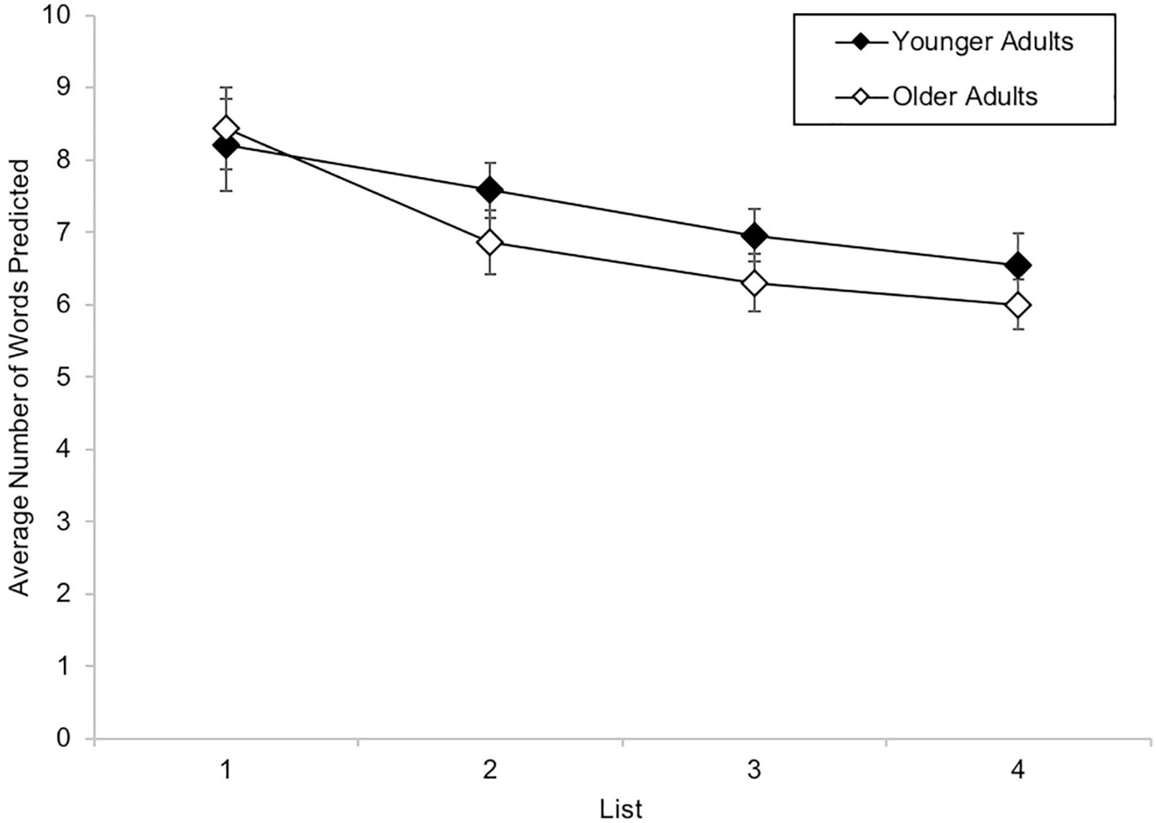

Figure 2.
Average number of words predicted for both age groups across lists in Experiment 1. Error bars represent ± 1 standard error of the mean.

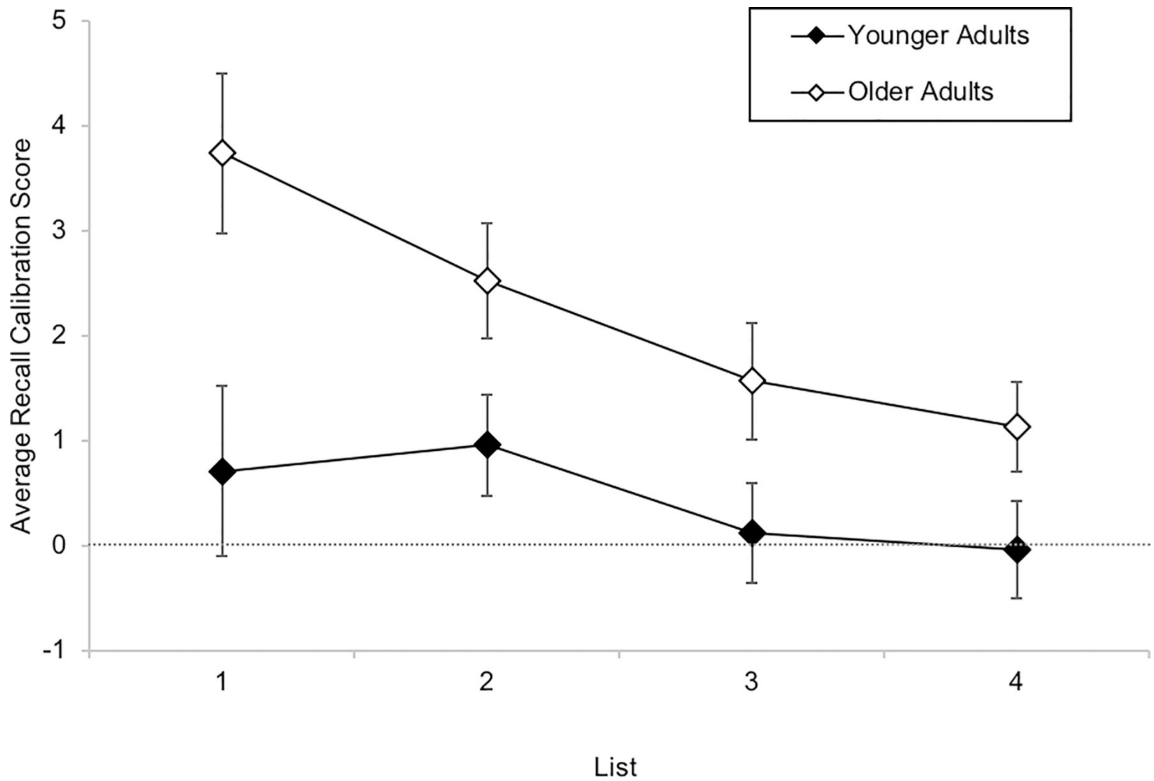

Figure 3.

Average recall prediction-performance difference scores in for both age groups across lists in Experiment 1. Dotted line indicates ideal accuracy score of zero. Error bars represent ± 1 standard error of the mean.

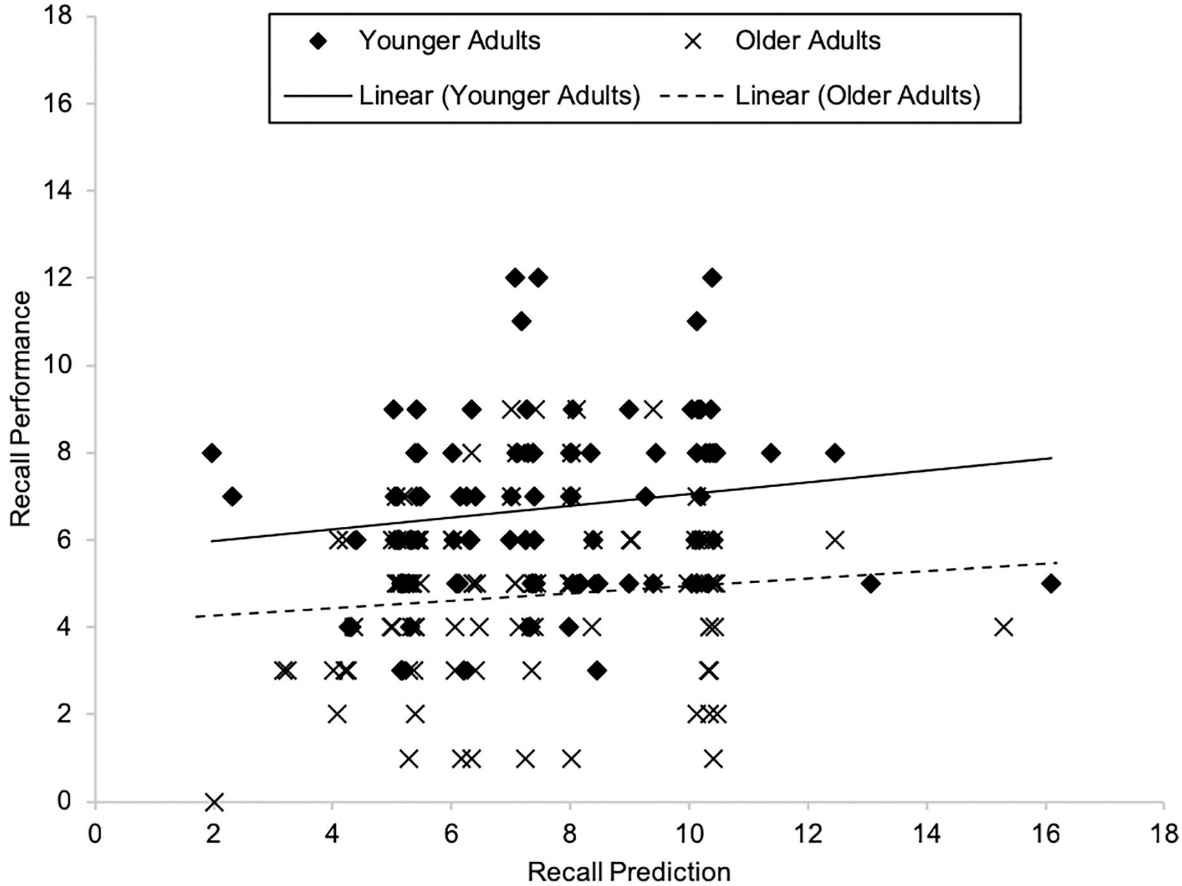

Figure 4.

Pearson's correlations for each age group between recall predictions and subsequent recall performance in Experiment 1 across all four lists. Trend lines indicate best linear fit. Points are slightly jittered along X-axis to minimize overlap where present.

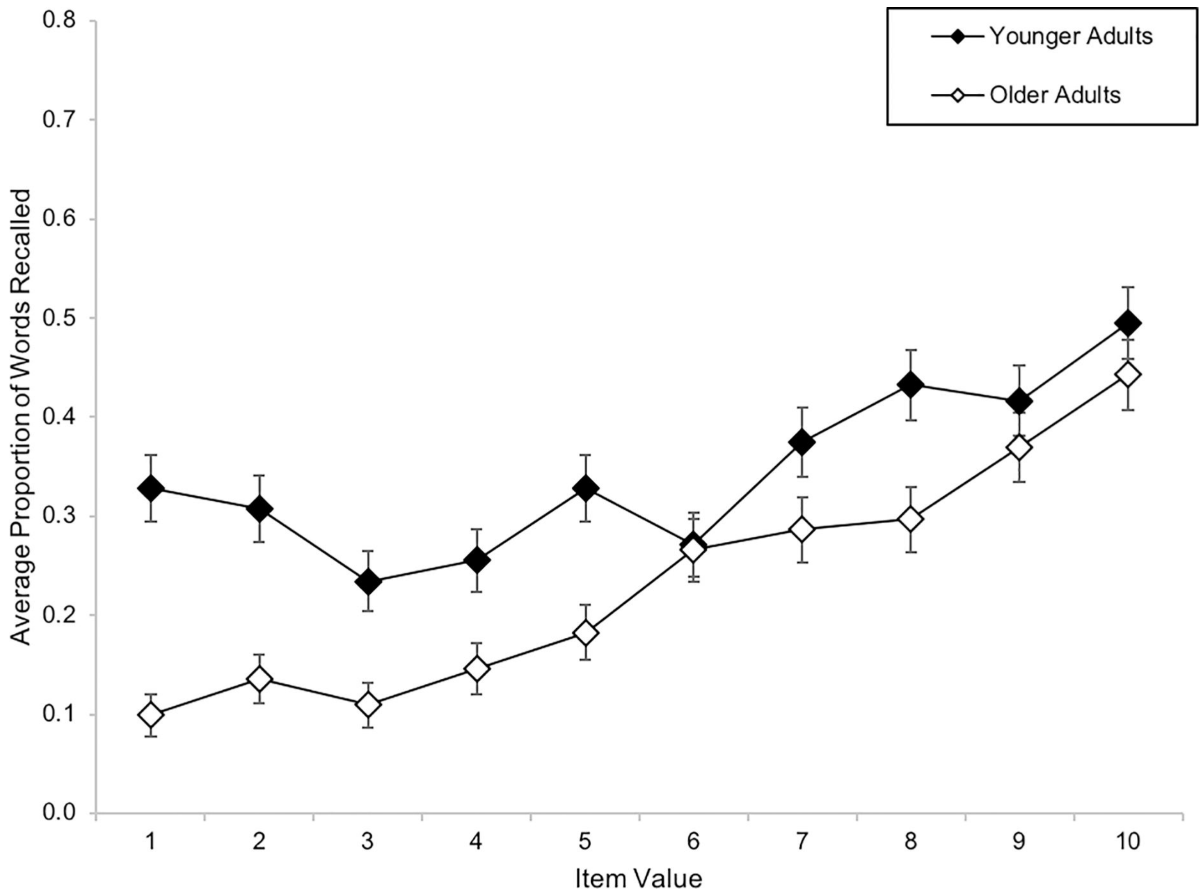

Figure 5.

Average proportion of words recalled as a function of age group and item value collapsed across lists in Experiment 1. Error bars represent ± 1 standard error of the mean.

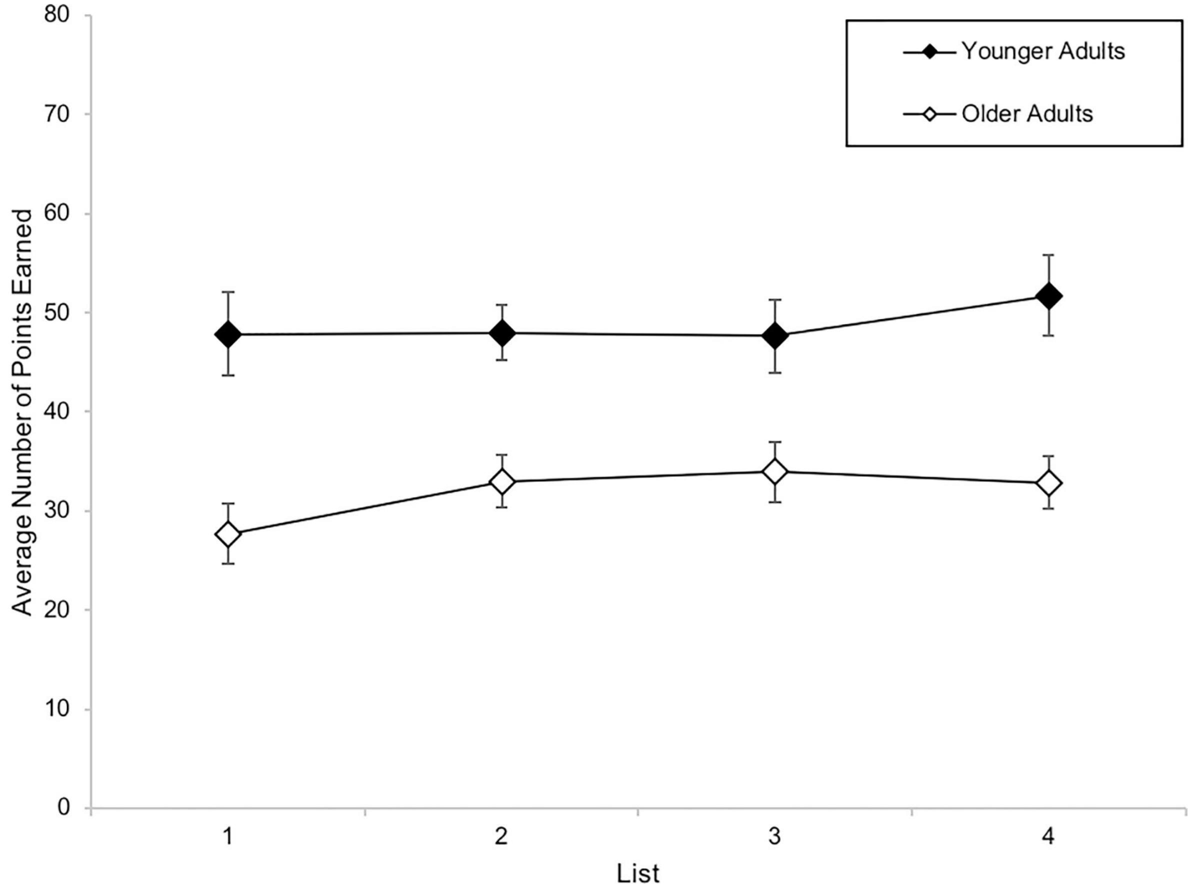

Figure 6.
Average number of points earned for both age groups across lists (out of 110 possible points) in Experiment 2. Error bars represent ± 1 standard error of the mean.

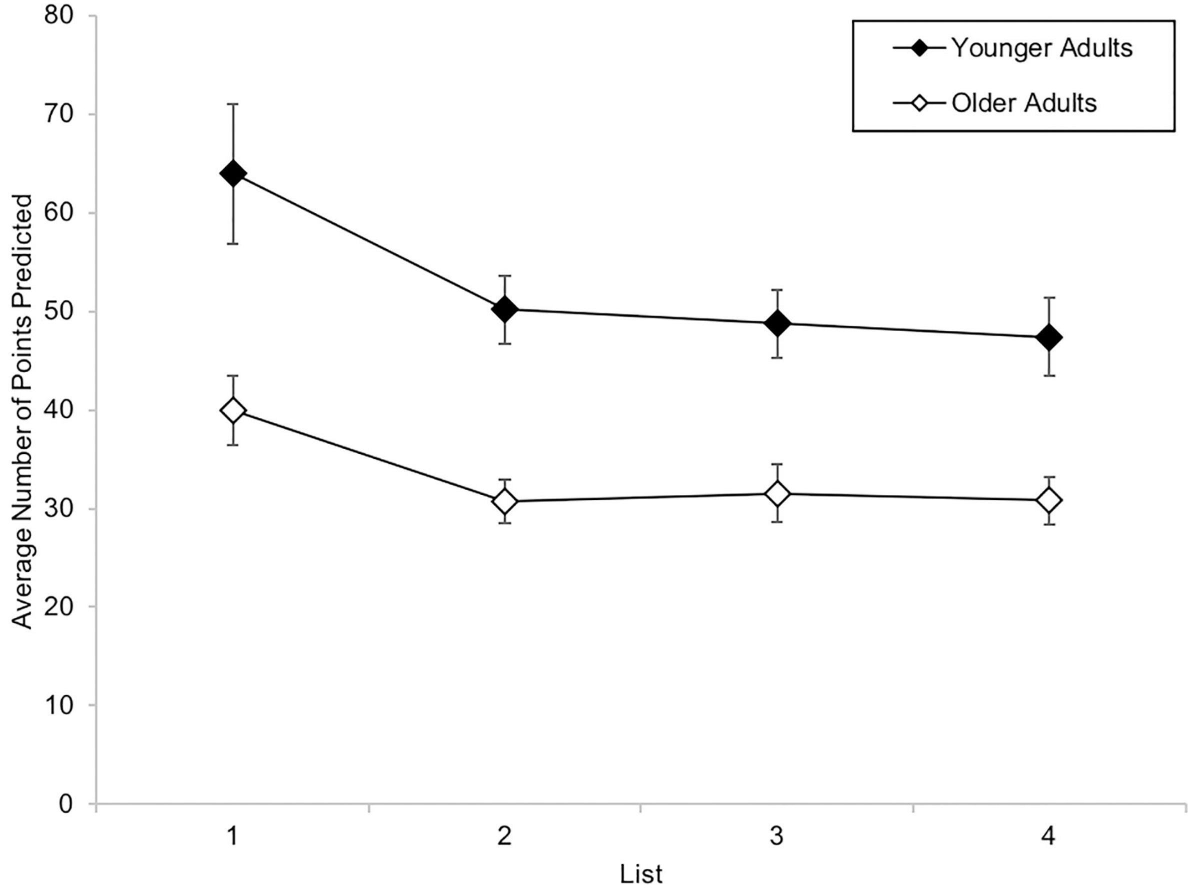

Figure 7.
Average number of points predicted across lists in Experiment 2. Error bars represent ± 1 standard error of the mean.

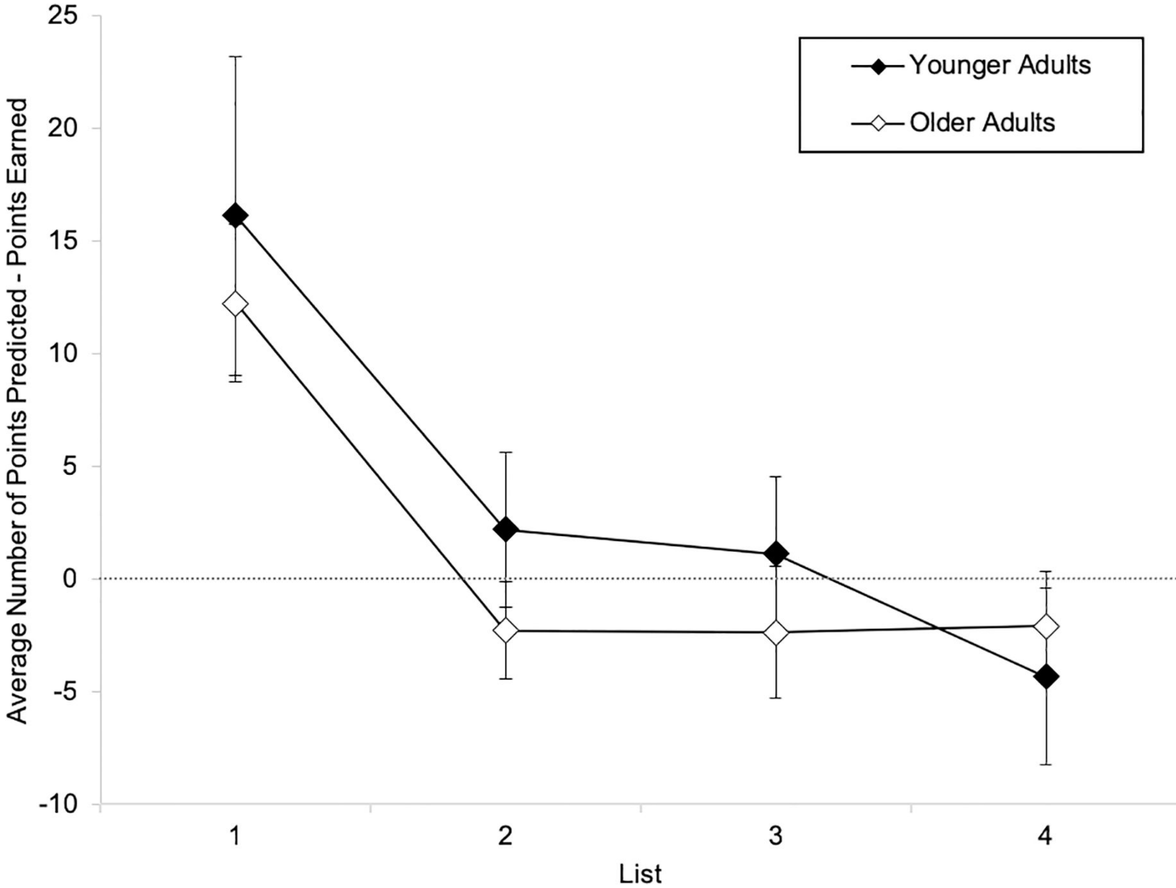

Figure 8.

Average points prediction-performance difference scores for both age groups across lists in Experiment 2. Dotted line indicates ideal accuracy score. Error bars represent ± 1 standard error of the mean.

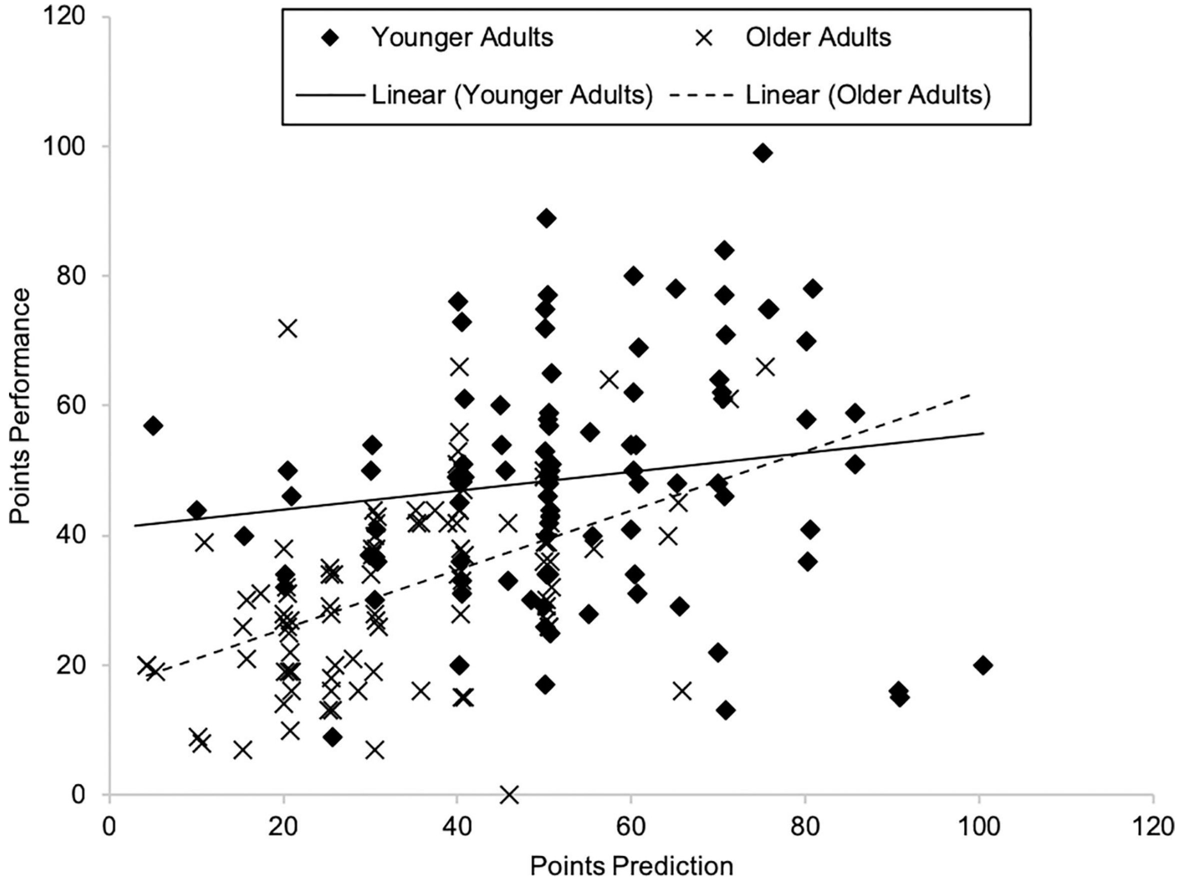

Figure 9.

Pearson's correlations for each age group between points predictions and subsequent points performance in Experiment 2 across all four lists. Trend lines indicate best linear fit. Points are slightly jittered along X-axis to minimize overlap where present.

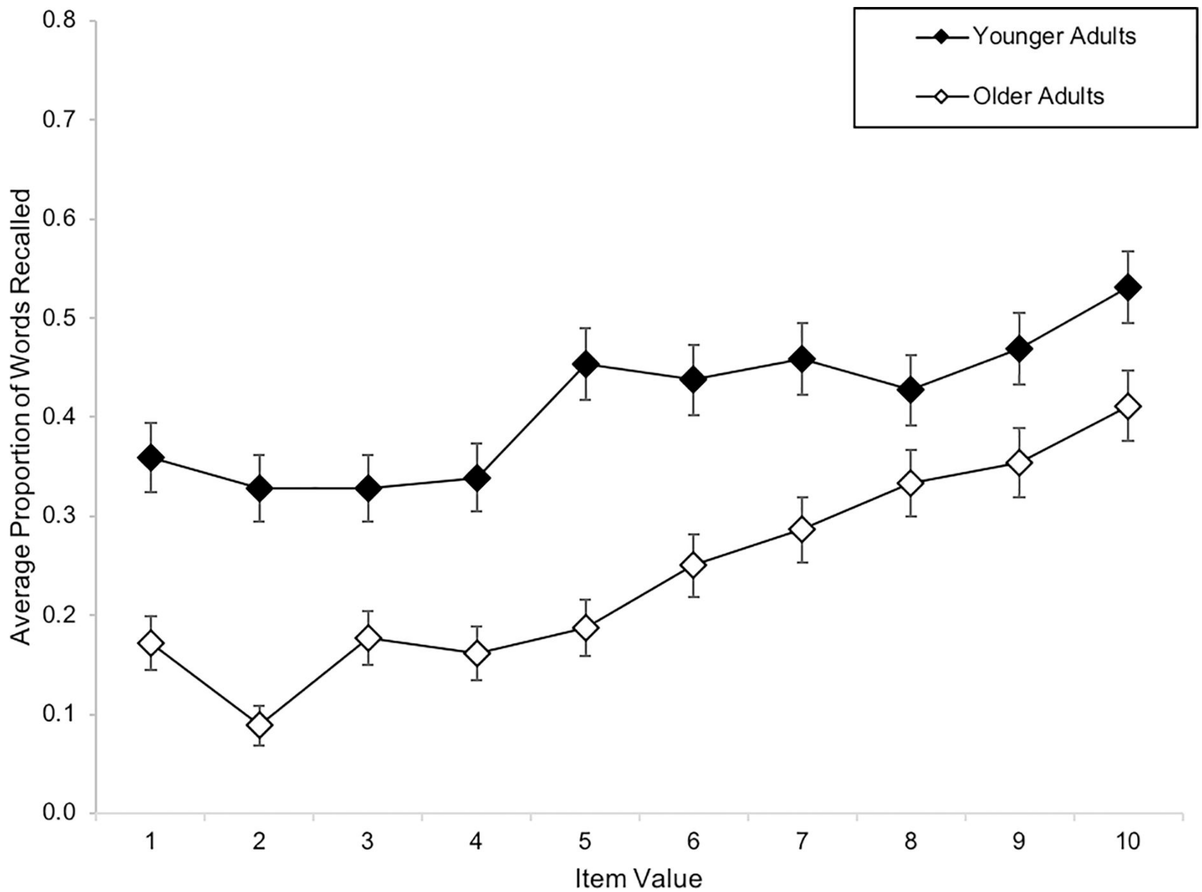

Figure 10.
Average proportion of words recalled as a function of age group and item value collapsed across lists in Experiment 2. Error bars represent ± 1 standard error of the mean.