Published in final edited form as:
Psychol Aging. 2008 December ; 23(4): 754–764. doi:10.1037/a0014295.

Targeting latent function: Encouraging effective encoding for successful memory training and transfer

Cindy Lustig and **Kristin E. Flegal**
Department of Psychology, University of Michigan

Abstract

Cognitive training programs for older adults often result in improvements at the group level. However, there are typically large age and individual differences in the size of training benefits. These differences may be related to the degree to which participants implement the processes targeted by the training program. To test this possibility, we tested older adults in a memory-training procedure either under specific strategy instructions designed to encourage semantic, integrative encoding, or in a condition that encouraged time and attention to encoding but allowed participants to choose their own strategy. Both conditions improved the performance of old-old adults relative to an earlier study (Bissig & Lustig, 2007) and reduced self-reports of everyday memory errors. Performance in the strategy-instruction group was related to pre-existing ability, performance in the strategy-choice group was not. The strategy-choice group performed better on a laboratory transfer test of recognition memory, and training performance was correlated with reduced everyday memory errors. Training programs that target latent but inefficiently-used abilities while allowing flexibility in bringing those abilities to bear may best promote effective training and transfer.

Keywords

AGING; MEMORY; TRAINING; TRANSFER; COGNITIVE REHABILITATION

Cognitive training programs provide a humbling – and often frustrating – reminder of how imprecise attempts to understand and modify human behavior remain. Training programs often fall into one of two categories: One method is to train participants on a specific strategy, such as the method of loci or the face-name mnemonic (e.g., Rebok & Balcerak, 1989; Yesavage & Rose, 1984). These programs often result in benefits on the training task but little or no transfer to other tasks. The other method uses a complex task (e.g., n-back, Jaeggi, Buschkuehl, Jonides, & Perrig, 2008) or set of tasks (e.g., Ball et al., 2002; Buschkuehl et al., this volume; Calero & Navarro, 2006; Craik et al., 2007; Loewenstein, Acevedo, Czaja, & Duara, 2004). This approach is more likely to show transfer to other tasks, but it is unclear what feature (s) of the training program are driving those gains. Another problem with many training programs is that the individuals who need training the most typically benefit the least: Both advanced age and lower initial ability are associated with reduced training benefits (Verhaeghen, Marcoen, & Goossens, 1992; Yesavage, Sheikh, Friedman, & Tanke, 1990).

Email correspondence should be addressed to Cindy Lustig (clustig@umich.edu) or Kristin Flegal (kflegal@umich.edu). Print correspondence should be addressed to Cindy Lustig, Psychology, University of Michigan, 530 Church Street, Ann Arbor, MI 48109-1043.

This paper reports the initial results from an ongoing project that analyzes age and individual differences in training and transfer with the goal of understanding the processes that underlie their success. As we will describe below, these analyses converge with longstanding theories of cognitive aging and relatively new neuroimaging data to suggest that one effective method for training older adults' memory is to encourage meaning-based, integrative processing (a.k.a. "deep" processing, Craik & Lockhart, 1972). However, allowing individuals to choose the exact way in which they implement this processing may be more effective than enforcing specific strategies. More generally, we hypothesize that successful programs are those that build on processes that remain functional in older adults, but that they often fail to use efficiently.

Our basic training procedure is a modified version of the repetition-lag procedure developed by Jennings and Jacoby (2003). This program was attractive because it was based on empirical and theoretical work on age differences in controlled, recollective retrieval processes (e.g., Jennings & Jacoby, 1993), was relatively easy to implement, and had already shown promise in positive transfer to laboratory tasks (Jennings, Webster, Kleyklamp, & Dagenbach, 2005; Jennings et al., 2006; see also the paper by Jennings in this issue). In each session, participants first learn 30 individually-presented words. This is followed by a yes/no recognition test that includes the 30 studied words as well as 30 unstudied words.

Critically, the unstudied words are repeated within the test list, so that when participants are faced with a familiar word, they must determine whether this familiarity arises because the word was on the study list or because it was a previously-presented lure. The demands on the controlled, recollective memory processes needed to make this discrimination are increased as the participant progresses through the training program. At first, only one and two items are interspersed between repetitions, making it relatively easy to recognize the repeated item. As soon as the participant reaches near-perfect performance at this lag level, the distance between items is increased to one and three items. Once near-perfect performance is achieved at this level, the lag interval is increased to two and four items, then to two and eight items, and so on.

The original procedure (Jennings & Jacoby, 2003; Jennings et al., 2005, 2006, this issue) restricts encoding time to two seconds per item. Modifying the program so that both encoding and retrieval were self-paced (Bissig & Lustig, 2007) revealed an important source of variability: The participants who performed well were those who spent proportionally more time on encoding than on retrieval. In fact, young adults spent nearly twice as long on the encoding phase for each word as did older adults (11.85 s vs 5.71 s), a rare example of longer response times being associated with better performance and younger age. Strikingly, statistically controlling for proportional encoding time eliminated the otherwise large performance differences related to age and individual differences in verbal ability.

Debriefing our participants suggested that it was not time *per se* that made the difference, but rather how individuals used that time. Those who performed best reported using strategies such as "Relate words to myself, and sometimes to each other", "Crafted stories with the words", or "Some words combined in sentences". Those who performed poorly reported either no conscious strategy or simply repeating the words to themselves. In keeping with their performance on the training task, "young-old" participants (75 years and less) more often reported meaning-based strategies, whereas participants of advanced age (76 years or more) more often reported no strategy or rehearsal-based strategies (Bissig & Lustig, 2007).

These reports fit nicely with Craik's classic framework for explaining age differences in memory (Craik & Byrd, 1982; Craik & Lockhart, 1972). This framework describes "deep", or meaning-based, relational processing at encoding as essential for supporting later episodic retrieval. Such processing is effortful and attention-demanding, and older adults often fail to self-initiate it in open-ended situations. In a recent expansion of this idea, Braver and colleagues have suggested that older adults tend to shift from *proactive* control – keeping a goal in mind and actively preparing for its execution – to *reactive* control – responding "on the fly" to a stimulus once it is presented and imminently demanding a response (Braver, Gray, & Burgess, 2007). Important from a training perspective, both of these frameworks emphasize that while older adults often fail to engage effortful, attention-demanding processes in open-ended situations, they can engage such processes when the task provides sufficient environmental support or constraint.

Neuroimaging data support this perspective and provide further insights. The idea that older adults frequently fail to self-initiate successful encoding processes receives support from the finding that when given intentional but otherwise unstructured encoding instructions ("Remember the words for a later memory test"), they fail to activate prefrontal cortex regions, particularly in left inferior frontal gyrus, associated with subsequent memory (e.g., Grady, et al., 1995). However, enforcing meaning-based encoding by requiring semantic judgments for each item brings older adults' activation in these regions to the level of young adults' (e.g., Logan, Sanders, Snyder, Morris, & Buckner, 2002; see also Lustig, et al., 2003; Lustig & Buckner, 2004). Consistent with the idea that these operations may be more difficult for older adults, they also frequently activate additional regions, with those older individuals who show more of this additional activation typically being the ones who show better memory performance (e.g., Cabeza, Anderson, Locantore, & McIntosh, 2002; Gutchess et al., 2005; see reviews by Cabeza, 2002; Reuter-Lorenz & Lustig, 2005).

Likewise, temporal differences in brain activation are consistent with an age-related shift to reactive control. For example, Head, Lustig, Isom, and Buckner (2006) found that young adults activated left inferior frontal gyrus more than did older adults during an intentional encoding task, but that older adults activated this region more during the retrieval test. This suggests that older adults may have failed to engage controlled, effortful encoding processes at the encoding stage, and then tried to compensate for this failure by increased activation of cognitive control when confronted with items requiring a response at retrieval. Even at the retrieval stage, older adults may have later or more temporally-extended prefrontal responses, again suggesting later and compensatory processing (Velanova, Lustig, Jacoby, & Buckner, 2007). However, older adults and even early Alzheimer's patients show learning-related plasticity in these regions (Lustig & Buckner, 2004), and may learn to engage these regions in intentional encoding conditions after a relatively short period of training (Kirchhoff, Anderson, Barch, & Jacoby, 2007).

In summary, the deep encoding processes supported by left inferior frontal gyrus exhibit several characteristics of an attractive target for training. As a group, older adults often fail to show efficient engagement of this and other memory-related regions in open-ended encoding situations, but some do, and these individual differences in activation have been correlated with individual differences in memory performance. This brain region shows learning-related plasticity, and age differences in its activation – presumably reflecting age differences in the processes in which it is involved – can be experimentally manipulated by changing the encoding task. Following training, older adults increase activation in left inferior frontal gyrus and report increased use of the deep encoding processes associated with it even under open-ended, intentional learning conditions. Although deep encoding is likely not the only way to engage the brain regions involved in successful memory, it is one that has a proven track record of working in older adult samples. Furthermore, methods based on deep encoding build upon older adults' strengths in semantic memory and vocabulary knowledge, areas in which they frequently outperform young adults (Verhaeghen, 2003).

What is the best way to train older adults to engage these processes? On the one hand, age deficits in self-initiation and proactive control suggest that strong enforcement of strategies that engage these processes may be needed, at least at first. On the other hand, previous training studies suggest that training specific strategies can paradoxically exacerbate age and ability differences, and that improvements on the training task often fail to transfer (Verhaeghen et al., 1992; Verhaeghen & Marcoen, 1996; Yesavage et al., 1990). Here, we compare these two approaches using the modified repetition-lag memory training procedure used in our previous study (Bissig & Lustig, 2007) and described above. The first condition (Integrated Sentences) controls encoding time and strongly enforces meaning-based, integrative encoding of the study lists. The second condition (Strategy Choice) controls encoding time and encourages participants to use meaning-based processing, but does not give specific strategy instructions.

Below we describe initial analyses of the effects of these encoding manipulations on training-task performance (also comparing them to the results of our earlier study, which was entirely open-ended with regards to encoding time and strategy) and transfer. With regards to training task performance itself, one hypothesis was that the enforced-encoding Integrative Sentences strategy might lead to the best performance if older adults failed to self-initiate good encoding in either the Open-Ended (Bissig & Lustig, 2007) condition or the Strategy Choice (controlled encoding time but not strategy) condition. On the other hand, the specific strategy used in the Integrative Sentences condition might also amplify age and ability differences and show limited transfer (Verhaeghen et al., 1992; Verhaeghen & Marcoen, 1996; Yesavage et al., 1990). For both the Integrated Sentences and Strategy Choice conditions, transfer effects (if found) were expected to be limited to verbal memory, since that is the target of our training program.

An important caveat is that the results presented here are part of an ongoing project. In particular, we do not yet have data from one of our planned control conditions: an enforced-rehearsal condition designed to mimic the strategies reported by the least-successful participants in our earlier study and to suppress the deep encoding strategies reported by our most-successful ones. Instead, our analyses focus on how enforcing a specific strategy versus simply encouraging older adults to spend sufficient time and attention at encoding influence group and individual differences in training performance and transfer.

Method

Participants

Thirty-two healthy older adults were assigned to the Integrated Sentences or Strategy Choice conditions, with the groups matched as closely as possible in age and education (n = 16 per group; see Table 1 for demographics). All participants were screened for medical or psychological conditions that could influence performance, and had Mini Mental State Evaluation scores (MMSE; Folstein, Folstein & McHugh, 1975) scores above 24 (mean = 28.9).

Materials and Procedure

Each participant completed eight study visits scheduled over the course of three weeks. The first day included informed consent procedures, a health and demographics questionnaire, dementia screening measures (MMSE and Short Blessed Test; Katzman et al., 1983), the Extended Range Vocabulary Test (ERVT; Educational Testing Services, 1976), and the baseline (pre-test) administration of potential transfer tasks (see descriptions below). At the end of the first day's visit, participants were given a brief (5 items) practice with the training task to familiarize them with the encoding instructions and time constraints for this portion of the study. On each of the following seven visits, participants completed four consecutive study-test cycles of the training task, described below. A questionnaire about strategies used in the training task and the post-training assessment of the transfer tasks occurred immediately following the training cycles on the last day.

The materials and procedure for the training task were identical to those used in Bissig and Lustig (2007) with the following exceptions: First, encoding time for each item was set to 14 seconds (this duration determined by pilot testing with another group of older adults), instead of being self-paced. Second, participants were randomly assigned to one of two experimental conditions, to determine which of two sets of encoding instructions they would receive. In the Integrated Sentences condition, participants were instructed to make a sentence out of each word and (for all but the first word in the list) the word that had just preceded it. These sentences were recorded to ensure compliance and for later content analysis. In the Strategy Choice condition, participants were instructed to think about the meaning of each word presented during encoding, in any way they would like, but with no explicit strategy specified.

Study and test words were chosen from the English Lexicon Project (Balota et al., 2002) and had a mean length of 5.76 letters and mean frequency of 20,487 out of 131 million. Length and frequency were balanced across lists and across conditions (studied, unstudied-short-lag, unstudied-long-lag). Each word was presented in large (32 point Arial) black-on-white font in the center of a computer screen. E-prime software was used for stimulus presentation and response collection (via keypress).

Participants completed four study-test cycles during each day's training session. During the study phase of the cycle, participants studied 30 individually presented words for 14 seconds each. During the test phase, participants were given a self-paced old/new recognition test, requiring discrimination of studied words from unstudied lures (c.f., Jennings & Jacoby, 2003; Jennings et al., 2005, 2006, this issue). The unstudied lures were repeated within the test list. Each recognition test had 90 items, pseudorandomly intermixed: The 30 studied words, the "new" first presentations of 30 unstudied lures, and the "repeated" second presentations of those same unstudied lures. Participants pressed one key (the "/" key) to indicate that it was one of the 30 studied items, a different key (the "z" key) to indicate that it was not. Each response on the retrieval test was followed by a feedback screen indicating accuracy (correct or incorrect) and trial type (studied, new, or repeated).

The difficulty of the retrieval test was gradually increased by increasing the lag between lure repetitions. Participants started at an easy level, with half of the lures repeated after only one intervening word, and the other half repeated after two intervening words (i.e., lag level 1 and 2). If the participant achieved criterion performance on the long-lag items (in this case, those with 2 items between repetitions), the difficulty of the next test was increased by moving to the next pair of lag intervals (i.e., 1 and 3 items between repetitions). The possible lag intervals were 1 and 2, 1 and 3, 2 and 4, 2 and 8, 4 and 12, 4 and 16, 8 and 20, 8 and 24, 12 and 28, 12 and 32, 16 and 36, and 16 and 40. Thus, participants were always working at one relatively easy lag interval, and one that might be more challenging. The criterion for moving up to the next lag level was set at 96% correct rejections of long-lag repeated lures for levels up to 2 and 8 (i.e., the 4th level) and relaxed to 93% for higher levels. Once a participant reached the maximum level (lags 16 & 40), s/he continued working at that level for the remaining sessions.

The transfer tasks administered before the first training session and immediately following the last session included some measures hypothesized to show transfer effects because they also emphasized semantically-based and/or integrative processing, and others hypothesized to not show transfer effects because they did not. These latter measures were included to test the specificity of any transfer effects and help identify which processes were being trained (e.g., if the training task primarily improves the engagement of semantically-based integrative processing, then we should not see improvement on tests such as the self-ordered pointing test.

However, if it primarily trains the rejection of repeated but currently-incorrect items [c.f., Jennings & Jacoby, 2003], then there should be positive transfer to the self-ordered pointing test).

The transfer tests thought to involve verbal or integrative memory included a shopping-list memory task and a face-name recognition task. For the shopping-list task, participants viewed 15 individually-presented words representing items that might appear on a typical grocery list (e.g., potatoes, soup). This was immediately followed by a 45-item self-paced old/new recognition test for the 15 studied items and 30 items that were not on the shopping list. As in the training task, participants indicated whether or not the item was a member of the study list by making a keypress response. In the face-name memory test, participants first viewed 10 face-name pairs, individually presented and self-paced, followed by a recall test in which they were presented with each face and asked to recall the name that had been paired with it. If the participant could not recall the name, they were given the correct name and a lure name and asked to identify the correct one.

The transfer tests thought to not involve verbal or integrative memory included the Pattern Comparison Test (Salthouse & Babcock, 1991), the Trail-Making Test (Armitage, 1945), and pattern and word versions of a self-ordered pointing test (SOPT; Attneave & Armoult, 1956). The Pattern Comparison test is a common measure of cognitive speed, and Version A of the Trail-Making Test is also thought to measure this construct. The SOPT and Version B of the Trail-Making Test are often used as measures of executive function. Our versions of the SOPT consisted of 16 words or 16 patterns arranged in a 4 × 4 grid. There were 16 pages for each test, and the 16 items for that test were arranged differently on each page. On each page, the task was to point to an item that had not been previously pointed to.

One danger with any intervention is that any benefits may influenced by placebo effects or factors not directly related to the target intervention (e.g., reporting improved memory because one thinks that memory training “should” improve memory). To assess this possibility, we also asked participants to complete the Memory Self Efficacy Questionnaire (MSEQ-4; Berry, West, & Dennehy, 1989) before training started and again after the last session. This questionnaire asks participants to rate their confidence in performing different memory tasks (e.g., remembering parts of a story or items on a shopping list) at different levels of difficulty (two items, eight items, etc.).

A word and source memory task was administered only at the end of training (i.e., no pre-test) because of concerns that prior exposure could dramatically change how participants approached the task, in particular how much attention they paid to item versus source information. Participants listened to 30 auditorily-presented words, half in a male voice and half in a female voice, randomly intermixed. This was followed by a visually-presented 60-item old-new recognition test. Participants indicated whether items were old or new, and if old, whether they had been spoken in a male or female voice. Two other tests (surprise recognition test for unstudied lures in the final training session and a false memory test) were included for comparison with the planned Rehearsal control group and will not be discussed in this paper.

Items were carefully screened across all transfer tasks to avoid overlap with the training procedure and with other transfer tasks. Also, with the exception of the Trail-Making Test, forms for the first- and last-day transfer tasks were not repeated (e.g., the 16 words and patterns used in the SOPT on Day 1 were a different set than those used on Day 8).

To assess any changes in real-world memory, we asked participants to complete the 35-item Everyday Memory Questionnaire (EMQ; Sunderland et al., 1983) on the first visit and before the training session on all following visits. (They could also opt to complete that day’s questionnaire before arriving in the lab.) Participants were asked to indicate how many times within the last 24 hours they had committed each of the memory errors listed on the questionnaire. The items can be split into five different subscales: Speech (e.g., “Finding that a word is on the tip of your tongue”), Reading/Writing (e.g., “Forgetting what the sentence you have just read was about and having to re-read it”), Faces/Places (e.g., “Failing to recognise television characters or other famous people by sight.”), Actions (e.g., “Discovering that you have done some routine thing twice by mistake.”), and New Things (e.g., “Forgetting to keep an appointment.”). The Speech and Reading/Writing subscales most obviously involve the processes that are the target of our training procedure.

Results

Effects of the enforced-encoding manipulation on training performance

We first asked how the encoding manipulations influenced progress on the training task itself, especially for “old-old” (age 76+ yrs) adults. The encoding groups did not differ in verbal ability (ERVT), speed (Pattern Comparison), years of education, or dementia scales (Table 1; all $p > .30$ for main effects of Encoding Group [Open-Ended, Integrative Sentences, Strategy Choice] and its interactions with Age Group [young-old, old-old]). However, training task performance as measured by the maximum lag at which participants reached criterion performance was particularly improved for old-old adults in the two enforced encoding-time conditions, $F(2, 45) = 5.99$, $p < .01$, such that age no longer predicted performance in these conditions (Figure 1).

We ranked training-task performance (1 = best, 16 = worst) in each of the two enforced encoding-time conditions using the same schema as in Bissig and Lustig (2007): Participants were ranked according to the highest lag level at which they achieved criterion performance, with ties between participants who reached the maximum lag level resolved by giving the better rank to the participant who reached criterion first (e.g., a participant who achieved this criterion in session 15 was assigned a better (lower) rank than one who reached it in session 21). Remaining ties were broken by assigning the better rank to the subject with better overall correct rejection of repeated lures.

Analyses with the ranking variable suggested that while enforcing encoding time eliminated age differences, enforcing a specific strategy amplified ability differences. (Figure 2, rows 1–3) In contrast with Bissig and Lustig (2007), age did not predict rank for either of the enforced encoding-time conditions used here, both $p > .30$. However, ability measures were even more strongly correlated with rank for the Integrated Sentences condition than they had been in the Open-Ended condition used in our earlier study. In contrast, ability measures did not predict training rank for the Strategy Choice condition, when encoding time was enforced but participants chose their own strategy. The most obvious explanation for these patterns is that enforcing encoding time encouraged participants to engage encoding processes they might not have otherwise brought to bear, but that their ability to adhere to the experimenter-instructed strategy (integrative sentences) was constrained by their pre-existing ability.

Contributions at retrieval

Although our manipulation emphasized encoding, there were also important individual differences and session-related changes (from the first to the last day) in how participants approached different item types on the retrieval test. The time for correct rejections of new vs repeated items did not differ by group or day and did not correlate with rank for either enforced encoding-time group (all $p > .10$), consistent with our previous analyses for the Open-Ended condition (Bissig & Lustig, 2007). Instead, and also in line with our previous analyses, the important differences were found in the time participants took to reject short- versus long-lag repeated items. The difference between short- and long-lag correct rejections was greater on the first day of training (1761ms vs 2073ms) than on the last day (1497ms vs 1548ms), F(1,30) = 13.11, p < .005; this effect did not interact with encoding group, F(1,30) = 1.39, p = .25. In other words, training was associated with an increased efficiency in rejecting long-lag items. Importantly, this did not simply reflect a bias to classify items as unstudied: Better rank was correlated with better accuracy on studied items, although this effect was only statistically significant for the Strategy Choice group (r = .68, p < .005; r = .43, p = .09 for the Integrated Sentences group)

Sensitivity to feedback was also related to performance. For both groups, participants with better ranks spent proportionally more time viewing feedback screens that followed incorrect responses than they did viewing feedback screens that followed correct responses, likewise replicating the results of our earlier study (Figure 2, row 4). The difference in time spent in looking at incorrect vs. correct feedback screens declined from the first to last day, F(1, 29) = 9.73, p < .005, but this effect did not interact with group, F(1, 29) = 2.03, p = .17. Correlations with rank fluctuated somewhat across days and groups but were generally positive (r = .27–.59). The patterns of longer times spent on incorrect than correct feedback were similar across item type (new, studied, short-lag, long-lag), F < 1. Correlations between rank and the proportion of time viewing responses might occur either because of a novelty effect (better performers saw incorrect screens less often), or because better performers were more likely to focus on negative feedback in an effort to avoid future mistakes. The current data cannot distinguish between these possibilities.

Transfer of training benefits to other tasks

We examined transfer both in terms of overall changes in scores from the first to the last day, and in how these changes correlated with performance in the training task itself (Table 2; Figure 3). As in our analyses of the training task itself, we used proportional scores wherever possible to reduce the influence of participants with extreme baseline scores, especially very fast or very slow RTs.

Contrary to our predictions, neither the shopping-list task nor the face-name memory task showed significant transfer effects. For the shopping-list task, this was likely due to ceiling effects; most participants had near-perfect performance even on the pre-test. For the face-name test, performance on the recall portion was very low on both pre- and post-test, but performance on the recognition portion was at ceiling. Discussions with our experimenters revealed that most of our participants were initially not producing responses for the recall question but were instead “holding out” for the easier two-choice recognition test on each face, upon which they performed at ceiling.

The only laboratory performance test to show robust pre- to post-test changes for both groups was the Trails test, in particular the more demanding Trails B portion that requires switching between letters and numbers. The Training Day (First, Last) X Form (A, B) interaction was significant, F(1, 30) = 5.02, p < .05, and did not interact further with Encoding Group, F(1, 30) = 1.71, p = .22. Post-hoc t-tests showed that speedups on Trails A were only marginal (41 s to 37 s, t(31) = 1.71, p = .10), corresponding with the lack of change on Pattern Comparison, another test of perceptual speed (9.9 items completed on both days, t < 1). In contrast, performance on Trails B improved from pre-to post-test (115 s vs 91 s, t(31) = 2.53, p < .05).

There are several possible interpretations of the changes in Trails B performance. One is that the training program generally improves controlled or executive processing, consistent with the idea that recollection training may broadly enhance these prefrontal-cortex dependent functions. In support of that idea, Jennings et al. (2006; see also their paper in this issue) found that training on their version of the procedure led to improvements in performance on the SOPT, n-back working memory task, and Digit Symbol Substitution test, all of which may have some working memory component. This possibility also receives indirect support from a recent paper by Persson and Reuter-Lorenz (2008) demonstrating that training on one set of tasks that reliably activate left inferior frontal gyrus results in relatively specific benefits to other tasks that also activate this region. Of interest, their study focused on tasks with an interference or conflict component like that found in the Trails B form but not on Trails A.

A less interesting possibility is that participants simply remembered the specific items they had seen on the first day when completing the tests again on the last day. Consistent with this possibility, participants in the current study did not show any pre/post training improvements in SOPT performance or the Pattern Comparison test. One difference between this study and those of Jennings et al. (2007; this issue) is that we used alternate forms of the SOPT and Pattern Comparison tests on the first and last day (i.e., different words or patterns on each day), whereas Jennings et al. used the same forms on both occasions. Therefore, it is possible that the improvements seen in their study resulted from item-specific practice effects.

The most intriguing possibility lies between these two extremes: Improvements on the Trails B test may have been to some degree dependent on the repetition of and memory for specific items, but the training program may have enhanced participants' ability to take advantage of this repetition. Jennings et al. (2005) found smaller transfer effects for a recognition-practice training group than for their recollection-training group, and smaller effects still for a no-contact control group that also repeated the tests within the same time window. This pattern is generally consistent with the idea that when effective, training programs may work by helping older adults make use of underlying, otherwise latent abilities and plasticity.

The Strategy Choice group performed significantly better than the Integrated Sentences group on item recognition (word memory) in the word and source memory task, $t(28) = 2.09$, $p < .05$, though not on source memory, $t(28) = 1.01$, $p > .30$. The differences in word memory may have arisen if individuals in the Strategy Choice group were more likely to apply the self-generated strategy they adopted during training to the word memory task: For individuals in this group, word memory correlated with training rank ($r = .55$, $p < .05$), but not with source memory ($r = .10$) or vocabulary ($r = .31$, $p =.25$). For those in the Integrated Sentences group, word memory appeared to be more strongly related to underlying ability, as its correlations with rank were only marginal ($r = .44$, $p = .10$), whereas it was significantly related to both source memory ($r = .62$, $p < .05$) and vocabulary ($r = .68$, $p < .01$). Partialling out vocabulary score eliminated the relationship between word and source memory, $r = .08$, further supporting the idea that underlying ability was an important factor in the Integrated Sentences condition.

Effects on self-report measures

Training was associated with a reduction in self-reported memory errors (Day 1 vs Day 8) as measured by the Everyday Memory Questionnaire (EMQ), $F(1,30) = 12.19$, $p < .005$, with no differences between the groups, $F < 1$. One concern is that since the EMQ is a self-report measure, the drop in memory errors could reflect a placebo effect: Participants might be reporting fewer memory errors because they *thought* they should be doing better as a result of training, not because they actually were experiencing fewer errors. We cannot definitively rule out this explanation without the Rehearsal control group, but several aspects of the data argue against this explanation.

First, scores on the Memory Self Efficacy Questionnaire (MSEQ) did not improve but instead numerically decreased from pre- to post-testing (Table 2). Second, EMQ errors showed an orderly decrease across the training period, consistent with the idea of gradual, training-related improvements (Figure 4). Third, our training program targeted verbal memory, and EMQ improvements were specific to the Speech ($t(31) = 4.00$, $p < .0005$) and Read/Write ($t(31) = 2.92$, $p < .01$) subscales; the Faces/Places scale showed a marginal improvement ($t(31) = 1.79$, p = .08), and the other subscales did not approach significance. In other words, EMQ improvements were specific to verbal memory, the trained domain.

Finally, although the mean MSEQ and EMQ results were similar for the two training groups, the correlation patterns were quite different (Figure 3). For participants in the Integrated Sentences group, poor rank in the training task was related to a drop in memory self-efficacy, and not related to EMQ changes. In contrast, for the Strategy Choice group, poor rank on the training task was related to a large number of reported memory errors on the first administration of the EMQ, and a drop after training. This pattern might occur if difficulties mastering the Integrated Sentences strategy and repeated exposure to error feedback reduced the confidence of poor performers, although the general focus on verbal encoding may still have had benefits (as suggested by overall improvements on verbal subscales of the EMQ). For the Strategy Choice group, the demands to develop one’s own effective encoding strategy within the training task may have encouraged individuals with everyday memory difficulties to look for similar opportunities in other situations. This suggestion is admittedly speculative. However, it fits well with the data from the word-list memory task described above, and offers intriguing possibilities for how training regimens might encourage compliance and transfer.

Discussion

Our data so far point to several interesting conclusions, as well as raising challenging questions for future research. First, they strongly suggest that the age differences in training progress seen in our previous study – and possibly in many training studies – may have occurred because participants of more advanced age did not approach the training regimen in the optimal way. In particular, in the open-ended encoding condition used in our previous study (Bissig & Lustig, 2007), the oldest (and least successful) participants spent only a short time encoding each word. In the current study, the training procedure strongly encouraged participants to devote time and attention to encoding each item, and the performance of our oldest participants was much improved. These patterns are consistent with the idea that aging is associated with declines in self-initiated processing and proactive control, but that the appropriate environmental supports or task constraints can overcome these declines (Craik & Byrd, 1982; Braver et al., 2007).

However, our data also join with previous training studies to suggest that teaching specific strategies is not the optimal way to institute such task constraints, and can lead to diminishing returns. The Integrated Sentences condition eliminated age differences, but amplified differences related to education and verbal ability, and was related to a decrease in memory self-efficacy in those participants who did not perform well. In contrast, pre-existing ability and education did not relate to performance in the Strategy Choice condition. Furthermore, better performance by the Strategy Choice group in the word-memory task and the correlations between training performance and improvements on the Everyday Memory Questionnaire suggest that this group may have been more likely to transfer their training to other tasks both in and out of the lab.

The correlations between pre-existing ability and training improvements seen in the open-ended encoding condition used in our previous study and in the Integrated Sentences condition used here are commonly seen in training research. In a meta-analysis of other training studies and concentrated analysis of their own empirical data, Verhaeghen, Goosens, and Marcoen (1992) found that approximately 40–65% of older adults in any given study failed to apply the methods in which they had been instructed. Even within the older adult groups, increased age was associated with decreased compliance with the instructed learning method. This pattern, by which “the rich get richer, the poor stay poor” was described more formally as the “amplification model” (Verhaeghen & Marcoen, 1996). In other words, training programs often amplify rather than reduce pre-existing age and ability differences in performance (see also Salthouse, 2006).

Verhaeghen and Marcoen (1996) suggest that this pattern occurs in part because of interactions between ability and compliance: Those participants who can implement the instructed learning methods are more likely to do so, and to do so more effectively. Their analyses focused on training studies with explicit strategy instruction (e.g., intentional encoding, method of loci, visual imagery). In studies that use complex tasks and open-ended instructions (e.g., n-back, Jaeggi et al., 2008; interference resolution training, Persson & Reuter-Lorenz, 2008), training benefits and transfer may likewise depend on the degree to which participants self-initiate successful processing in the training task.

Consistent with this idea, Derwinger and colleagues reported that if older adults successfully self-generate strategies that improve their performance on a memory task they show longer-lasting improvements than those instructed on an experimenter-chosen mnemonic, and are more likely to continue using those strategies long after the formal training period has ended (Derwinger, Neely, & Backman, 2005; Derwinger, Neely, McDonald, & Backman, 2005). Our findings of better word-list memory and stronger correlations between training rank and EMQ improvement for the Strategy Choice group suggest that self-generated strategies may also be more likely to transfer.

This brings us to the most difficult question: What processes underlie training improvements and the apparent transfer effects seen in the current study and in studies (e.g., Jennings & Jacoby, 2003; Jennings et al., 2005, 2006) that restrict encoding time to short (2 s) periods? Jennings et al. emphasize the importance of retrieval processes, whereas our results put the emphasis on encoding. This may be due in part to procedural differences (2 s encoding versus unrestricted or 14 s encoding; the use of repeated versus alternate forms for transfer tests).

There were some effects at retrieval. In particular, participants became more efficient at rejecting long-lag items.[1] As described in Bissig and Lustig (2007), we believe that this occurs because improved encoding allows improved use of what Jacoby has termed *source-constrained retrieval*: A proactive retrieval mode by which participants try to restrict retrieval and memory access to the target source. In the words of our participants on the post-test questionnaire, “I tried to remember the sentences I used in the study portion”, “By ignoring the new words, the repetition didn’t bother me.” This form of retrieval is thought to be heavily dependent on the quality of encoding (Jacoby, Shimiuzu, Velanova, & Rhodes, 2005). However, source-constrained retrieval sometimes fails. When it does, participants must retroactively determine why an item seems familiar, and greater difficulty with this process (as reflected by proportionally longer RTs for long-lag repeated items) is associated with worse recollection. This *source identification* process is most likely the major locus of change in versions of the training program that restrict encoding processes (Jennings & Jacoby, 2003;Jennings et al., 2005,2006, this issue). In the current study, differences in long- vs short-lag item retrieval times did not correlate with training rank, did not differ by encoding group, and did not correlate with any of the transfer-task outcomes. By contrast, differential attention to feedback after incorrect versus correct items did not change significantly over the training period, but did correlate with training rank (see r of Figure 2). As mentioned earlier, it is hard to know from this dataset whether these correlations reflect greater self-initiation by good

---

[1] A reviewer (Gus Craik) suggested that simple improvements in encoding might be enough to lead to this effect, without a direct influence on retrieval processes. We tend to agree that the major action in this training procedure is at the encoding end. However, given that the procedure was originally designed to improve recollective processes (Jennings & Jacoby, 2003), it seemed important to give also fair consideration to potential changes at the retrieval end.

performers (greater attention to current mistakes in order to avoid future ones) or a simple novelty effect (since good performers would see incorrect feedback less often).

Our manipulations focused on the encoding stage, with effects both on the training task itself and on the transfer tasks. Both the controlled-strategy (Integrated Sentences) and Strategy Choice conditions showed apparent improvements in everyday memory. Furthermore, even though both groups trained on a verbal task, the Strategy Choice group showed better word memory and stronger correlations with improvements in the Everyday Memory Questionnaire. Both of these patterns provide support for the idea that self-generated strategies may be especially likely to encourage successful transfer.

The improvements on the Everyday Memory Questionnaire should be interpreted with some caution, given that it is a self-report measure and that we do not yet have control-group data. However, several points argue against the idea that improvements on the EMQ only reflect a placebo effect. First, if improvements on the EMQ reflected a bias to believe that one's memory was getting better as a result of training, then we should have also seen an increase in scores on the Memory Self-Efficacy Questionnaire (MSEQ). Instead, participants' scores on the MSEQ tended to *decrease* over the training period. It is difficult to reconcile this apparent decline in participants' confidence about their memory abilities with a placebo-effect explanation of the EMQ changes.

Second, if improvements on the EMQ reflected either a bias to think that one's memory was getting better as a result of training or some other effect due to simply completing the questionnaire on multiple occasions (e.g., better sensitivity to one's daily memory errors as a result of monitoring them), then we should have seen improvements on all of the subscales. Instead, the improvements were largely confined to those subscales that relate to verbal memory processes, the target of our training program. Finally the specific improvements for participants who originally showed the most memory errors and correlations with training rank in the Strategy Choice condition are also more consistent with a training effect.

In summary, our results suggest that targeting functions that are relatively preserved but inefficiently-used by older adults may be an effective method for promoting training and transfer. As described above, older adults can engage deep encoding processes – and in fact have better semantic knowledge than do young adults, providing an especially good basis for such processes – but often fail to self-initiate such processes in open-ended situations (Craik & Byrd, 1982, Braver et al., 2007). The appropriate environmental supports or task constraints can remediate these failures in self-initiation and proactive control; in our case these constraints were put in place by enforcing encoding time. On the other hand, *too* much constraint can have diminishing returns – restricting participants to a specific strategy tended to amplify pre-existing ability differences, consistent with previous work (Verhaeghen & Marcoen, 1996). The restricted-strategy condition also appeared to be less effective than the strategy-choice condition in encouraging transfer both to a laboratory memory test and reducing everyday memory errors.

It is doubtful that participants in either condition spent 14 seconds encoding every to-be-remembered item either in the laboratory transfer tests or in everyday life. Instead, the training procedure likely encouraged a more proactive approach to attention and deep processing at encoding.[2] We note with interest that Kirchhoff et al. (2007) found that after briefly training older adults on several different strategies designed to promote deep encoding, participants differed in the particular strategy they chose on a later open-ended memory test, but all were effective in improving memory performance and increasing frontal brain activations.

---

[2] We thank Morris Moscovitch for this question.

Our results join with theirs to suggest that manipulations that encourage older adults to increase proactive control and deep encoding but that allow flexibility in the method of doing so are the most likely to promote successful training and transfer. One caveat is that participants in both studies were relatively well-educated; participants of especially low education or ability may still benefit from greater guidance, at least initially. However, the improvements on laboratory memory tests found in both studies and the correlations with reduced real-world memory errors found in the current dataset are especially exciting, given that transfer effects are traditionally the Achilles' heel in training research.

We end our paper with a statement that will probably be found in all papers in this special issue: "More research is needed" to understand what factors underlie successful training and transfer to real-world tasks. Our data suggest that identifying remaining but inefficiently-used abilities in older adults, and finding ways to encourage the more optimal use of those remaining functions, will be an important part of that endeavor.

Acknowledgments

This work was funded by the National Institute on Aging (R03 AG029329-01A1 from the National Institute on Aging to Cindy Lustig, P30 AG024824 to the University of Michigan Older Americans Independence Center), by the Office of the Vice President of Research at the University of Michigan, and by a University of Michigan Rackham Spring/Summer Faculty Fellowship. We thank Ulrich Mayr for organizing this special issue on training, and are grateful to the members of the Lustig and Reuter-Lorenz laboratories for their comments and assistance throughout this project. Monica Kumar, Taryn Weil, Lauren Schiff, and Monika Potrykus were especially helpful with participant recruitment. Lauren Schiff and Mary Winters also assisted with references.

References

Armitage SG. An analysis of certain psychological tests used for the evaluation of brain injury. Psychological Monographs 1945;60:1–48.

Attneave F, Arnoult MD. Methodological considerations in the quantitative study of shape and pattern perception. Psychological Bulletin 1956;53:221–227.

Ball K, Berch DB, Helmers KF, Jobe JB, Leveck MD, Marsiske M, Morris JN, Rebok GW, Smith DM, Tennstedt SL, Unverzagt FW, Willis SL. Advanced Cognitive Training for Independent and Vital Elderly Study Group. Effects of cognitive training interventions with older adults: A randomized controlled trial. Journal of the American Medical Association 2002;288:2271–2281. [PubMed: 12425704]

Balota, DA.; Cortese, MJ.; Hutchison, KA.; Neely, JH.; Nelson, D.; Simpson, GB., et al. Washington University; 2002. The English lexicon project. A web-based repository of descriptive and behavioral measures for 40,481 English words and nonwords. http://elexicon.wustl.edu/

Berry JM, West RL, Dennehy D. Reliability and validity of the Memory Self-Efficacy Questionnaire (MSEQ). Developmental Psychology 1989;25:701–713.

Bissig A, Lustig C. Who benefits from psychological training? Psychological Science 2007;18:720–726. [PubMed: 17680944]

Braver, TS.; Gray, JR.; Burgess, GC. Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In: Conway, ARA.; Jarrold, C.; Kane, MJ.; Miyake, A.; Towse, JN., editors. Variation in Working Memory. Oxford University Press; 2007. p. 76-106.

Buschkuehl M, Jaeggi SM, Hutchison S, Perrig-Chiello P, Däpp C, Müller M, Breil F, Hoppeler H, Perrig WJ. Differential effects of cognitive and physical training on memory performance in old-old. (this issue)

Cabeza R. Hemispheric asymmetry in older adults: The HAROLD model. Psychology and Aging 2002;17:85–100. [PubMed: 11931290]

Calero MD, Navarro E. Cognitive plasticity as a modulating variable on the effects of memory training in elderly persons. National Academy of Neuropsychology 2006;22:63–72.

Craik, FIM.; Byrd, M. Aging and cognitive deficits: The role of attentional resources. In: Craik, FIM.; Trehub, S., editors. Aging, and cognitive processes. New York: Plenum Press; 1982. p. 191-211.

Craik FIM, Lockhart RS. Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior 1972;11:671–684.

Craik FIM, Winocur G, Palmer H, Binns MA, Edwards M, Bridges K, Glazer P, Chavannes R, Stuss DT. Cognitive rehabilitation in the elderly: Effects on memory. Journal of the International Neuropsycological Society 2007;13:132–142.

Cabeza R, Anderson ND, Locantore JK, McIntosh AR. Aging gracefully: Compensatory brain activity in high- performing older adults. Neuroimage 2002;17:1394–1402. [PubMed: 12414279]

Deese J. On the prediction of occurrence of particular verbal intrusions in immediate recall. Journal of Experimental Psychology 1959;58:17–22. [PubMed: 13664879]

Derwinger A, Neely A, Bäckman L. Design your own memory strategies! Self-generated strategy training versus mnemonic training in old age: An 8-month follow-up. Neuropsychological Rehabilitation 2005;15:37–54. [PubMed: 16353852]

Derwinger A, Neely A, MacDonald S, Bäckman L. Forgetting numbers in old age: Strategy and learning speed matter. Gerontology 2005;51:277–284. [PubMed: 15980656]

Educational Testing Services. Princeton, NJ: 1976. Kit of factor-referenced tests.

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

Grady CL, McIntosh AR, Horwitz B, Maisog JM, Ungerleider LG, Mentis MJ, Pietrini P, Schapiro MB, Haxby JV. Age-related reductions in human recognition memory due to impaired encoding. Science 1995;269:218–221. [PubMed: 7618082]

Gutchess AH, Welsch RC, Hedden T, Bangert A, Minear M, Liu LL, Park DC. Aging and the neural correlates of successful picture encoding: Frontal activations compensate for decreased medial-temporal activity. Journal of Cognitive Neuroscience 2005;17:84–96. [PubMed: 15701241]

Head, D.; Lustig, C.; Isom, M.; Buckner, RL. Age differences in frontal recruitment between encoding and retrieval; Atlanta, GA. Poster presented at the annual meeting of the Society for Neuroscience; 2006 Oct.

Jacoby LL, Shimizu Y, Daniels KA, Rhodes MG. Modes of cognitive control in recognition and source memory: Depth of retrieval. Psychonomic Bulletin & Review 2005;12:852–857. [PubMed: 16524001]

Jaeggi SM, Buschkuehl M, Jonides J, Perrig WJ. Improving fluid intelligence with training on working memory. Proceedings of the National Academy of Sciences USA 2008;105:6829–6833.

Jennings, JM.; Carello, EA.; Dagenbach, D.; Rapp, SR.; Brenes, GA.; Atkinson, HH. Improving recollection in individuals with mild cognitive impairment: Training and transfer effects. Atlanta, GA. Cognitive Aging Conference; 2006.

Jennings JM, Jacoby LL. Automatic versus intentional uses of memory: Aging, attention, and control. Psychology and Aging 1993;8:283–293. [PubMed: 8323731]

Jennings JM, Jacoby LL. Improving memory in older adults: Training recollection. Neuropsychological Rehabilitation 2003;13:417–440.

Jennings JM, Webster LM, Kleykamp BA, Dagenbach D. Recollection training and transfer effects in older adults: Successful use of a repetition-lag procedure. Aging, Neuropsychology, and Cognition 2005;12:278–298.

Katzman R, Brown T, Fuld P, Peck A, Schechter R, Schimmel H. Validation of a short orientation-memory-concentration test of cognitive impairment. American Journal of Psychiatry 1983;140:734–739. [PubMed: 6846631]

Kirchhoff, BA.; Anderson, BA.; Barch, DA.; Jacoby, LL. The effects of encoding strategy and cognitive training on age-related changes in prefrontal activity; San Diego, CA. Poster presented at the Society for Neuroscience Meeting; 2007 Nov.

Loewenstein DA, Acevedo A, Czaja SJ, Duara R. Cognitive rehabilitation of mildly impaired Alzheimer disease patients on cholinesterase inhibitors. American Journal of Geriatric Psychiatry 2004;12:395–402. [PubMed: 15249277]

Logan JM, Sanders AL, Snyder A, Morris JC, Buckner RL. Under-recruitment and non-selective recruitment: Dissociable neural mechanisms associated with aging. Neuron 2002;33:827–840. [PubMed: 11879658]

Lustig C, Buckner RL. Preserved neural correlates of priming in old age and dementia. Neuron 2004;42:865–875. [PubMed: 15182724]

Lustig C, Snyder AZ, Bhakta M, O'Brien KC, McAvoy M, Raichle ME, Morris JC, Buckner RL. Functional deactivations: Change with age and dementia of the Alzheimer type. Proceedings of the National Academy of Sciences USA 2003;100:14504–14509.

Persson, Reuter-Lorenz. Gaining control: Training of executive function and far transfer of the ability to resolve interference. Psychological Science 2008;19:881–889. [PubMed: 18947353]

Rebok GW, Balcerak LJ. Memory self-efficacy and performance differences in young and old adults: The effect of mnemonic training. Developmental Psychology 1989;25:714–721.

Reuter-Lorenz PA, Lustig C. Brain aging: reorganizing discoveries about the aging mind. Current Opinion in Neurobiology 2005;15(4):245–251. [PubMed: 15831410]

Roediger HL, McDermott KB. Creating false memories: Remembering words not presented in lists. Journal of Experimental Psychology: Learning, Memory, & Cognition 1995;21:803–814.

Salthouse TA. Mental exercise and mental aging: Evaluating the validity of the “use it or lose it” hypothesis. Perspectives on Psychological Science 2006;1:68–87.

Salthouse TA, Babcock RL. Decomposing adult age differences in working memory. Developmental Psychology 1991;27:763–776.

Sunderland A, Harris JE, Baddeley AD. Do laboratory test predict everyday memory? A neuropsychological study. Journal of Verbal Learning and Verbal Behavior 1983;22:341–357.

Velanova K, Lustig C, Jacoby LL, Buckner RL. Evidence for frontally-mediated controlled processing differences in older adults. Cerebral Cortex 2007;17:1033–1046. [PubMed: 16774962]

Verhaeghen P, Marcoen A. On the mechanisms of plasticity in young and older adults after instruction in the method of loci: Evidence for an amplification model. Psychology and Aging 1996;11:164–178. [PubMed: 8726382]

Verhaeghen P, Marcoen A, Goossens L. Improving memory performance in the aged through mnemonic training: A meta-analytic study. Psychology and Aging 1992;7:242–251. [PubMed: 1535198]

Verhaeghen P. Aging and vocabulary score: A meta-analysis. Psychology and Aging 2003;18(2):332–339. [PubMed: 12825780]

Yesavage JA, Rose TL. The effects of a face-name mnemonic in young, middleaged, and elderly adults. Experimental Aging Research 1984;10:55–57. [PubMed: 6734688]

Yesavage JA, Sheikh JI, Friedman L, Tanke E. Learning mnemonics: Roles of aging and subtle cognitive impairment. Psychology and Aging 1990;5:133–137. [PubMed: 2317292]

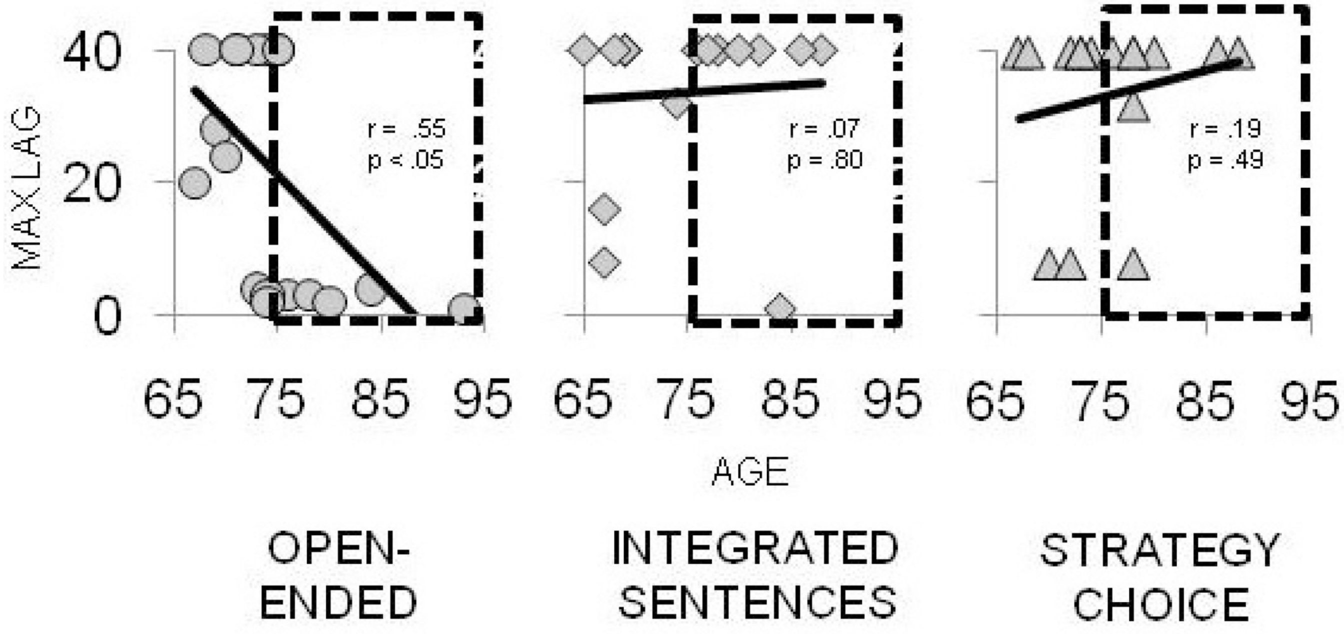

Figure 1.

Enforcing encoding time eliminates age differences in training performance. The y-axis shows the maximum lag at which participants achieved criterion performance; x-axis shows age. Boxes highlight “old-old” participants age 76+ yrs.

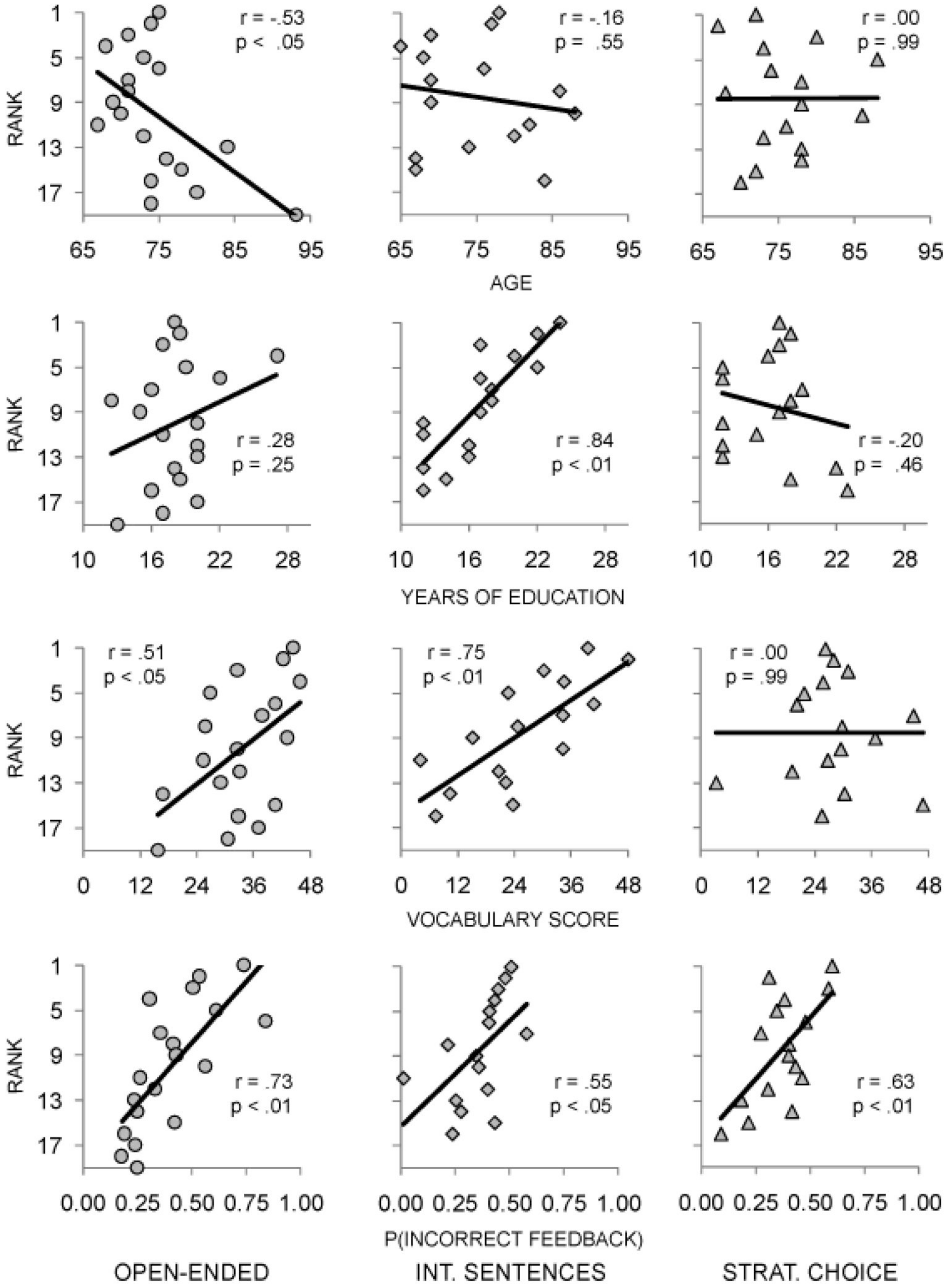

Figure 2.

Correlations between rank and age, ability measures, and attention to feedback after incorrect items. Education and verbal ability predict training rank in the Open-Ended and Integrative Sentences condition, but not in the Strategy Choice condition, which encouraged attention to encoding but did not enforce a specific strategy. Greater proportional time spent looking at feedback after incorrect items than after correct items (p(incorrect) = (incorrect feedback RT – correct feedback RT) / (incorrect feedback RT + correct feedback RT)) was related to good performance in the training task in all conditions.

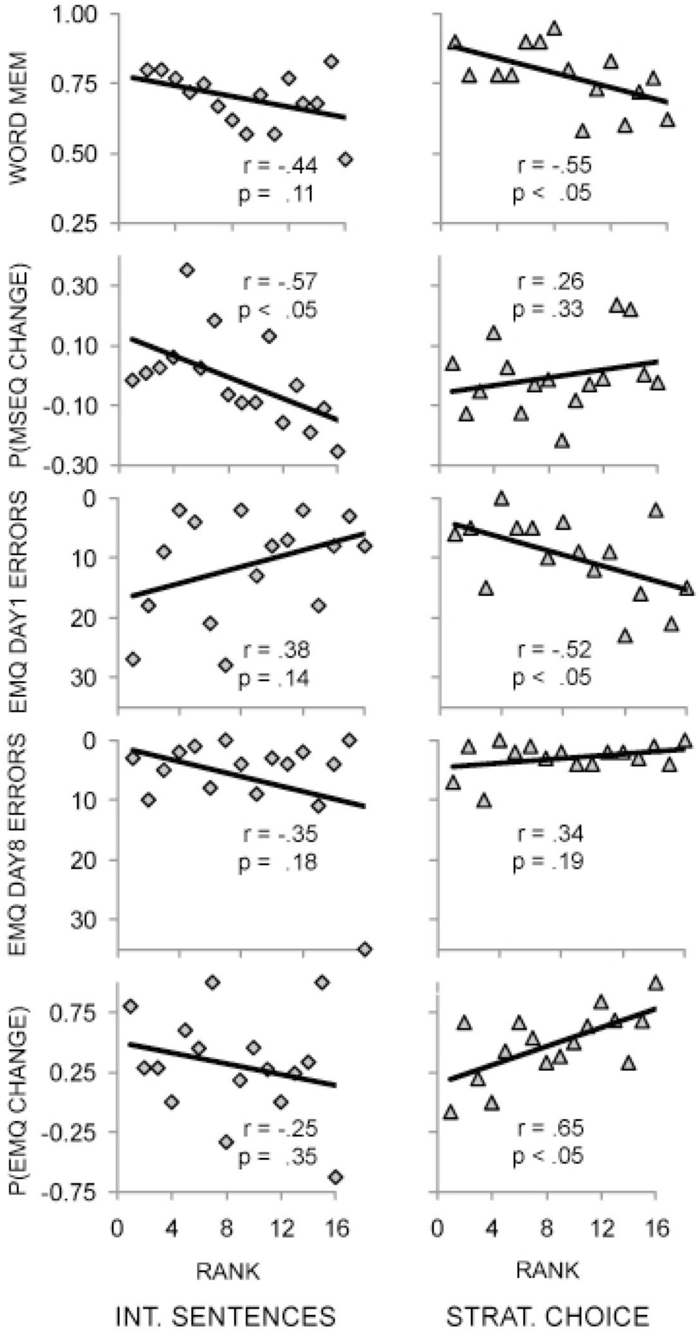

Figure 3.

Correlations between training rank and outcome measures. Axes are arranged so that better scores (e.g., higher self-efficacy, fewer memory errors) are always higher on the y-axis. Performance in the Strategy Choice condition correlates with both word memory and the Everyday Memory Questionnaire (EMQ). Removing the subject with the outlying score in the last administration of the EMQ changes the rank X EMQ8 correlation for the Integrated Sentences condition to r = .09, the rank X EMQ-change correlation to r = −.01.

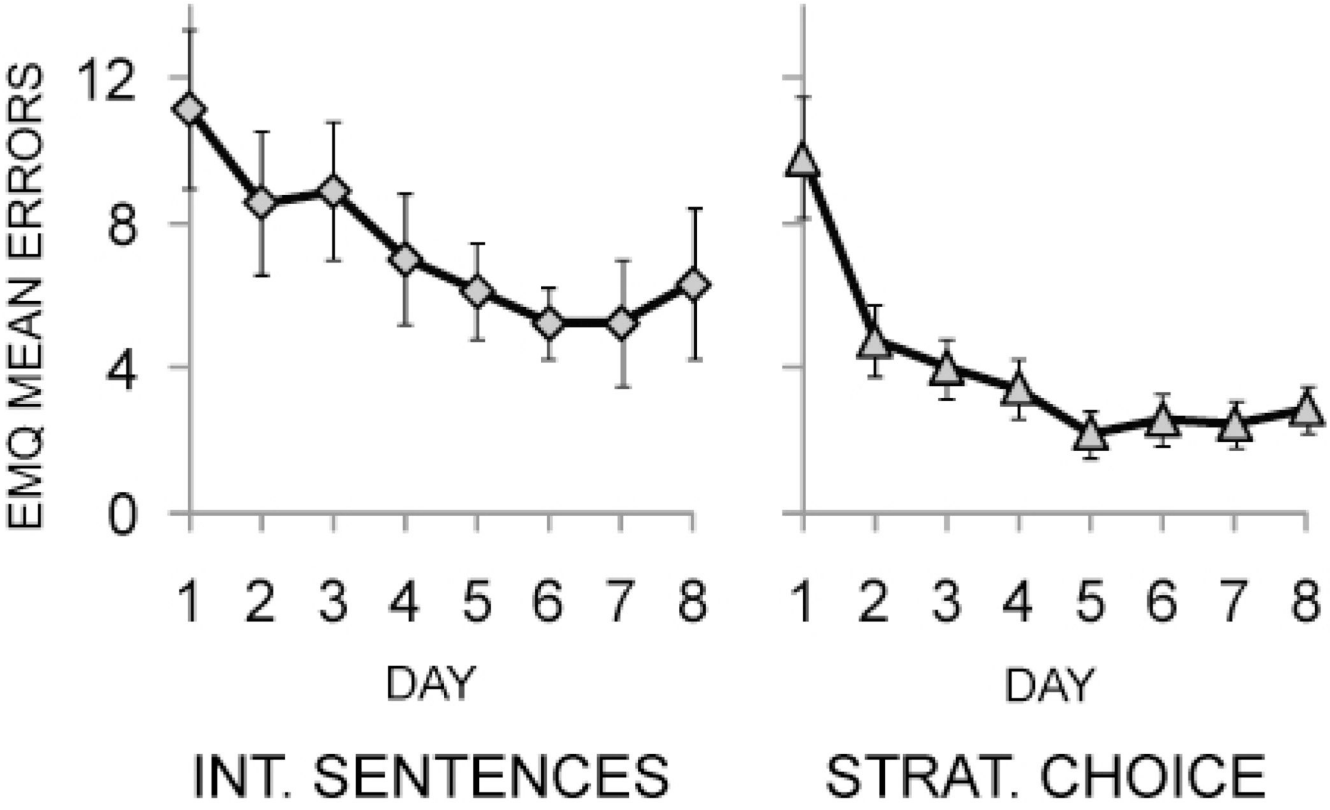

Figure 4.

Everyday memory errors show an orderly decline over the training period. Y-axis shows mean daily errors reported on the EMQ for each group.

Table 1

Participant demographics by condition

	Open-Ended (N = 19)		Integrated Sentences (N = 16)		Strategy Choice (N = 16)	
	M	SD	M	SD	M	*SD*
Age (years)	74.5	6.1	74.9	7.5	75.7	5.8
Education (years)	18.1	3.2	16.8	3.8	16.3	3.5
ERVT	33.2	8.7	25.8	12.5	27.8	10.1
MMSE	28.6	1.4	28.9	1.5	29.0	1.2
SBT	1.0	1.2	1.3	1.2	1.6	2.3

Note: ERVT = Extended Range Vocabulary Test (maximum score = 48); MMSE = Mini Mental State Evaluation (maximum score = 30; higher scores = better performance); SBT = Short Blessed Test (maximum score = 28; higher scores = worse performance)

Table 2

Transfer task performance by condition

	Integrated Sentences		Strategy Choice	
	Pre-training **MSD**	**Post-training** **MSD**	**Pre-training** **MSD**	**Post-training** **MSD**
Shopping List recognition accuracy	93.9%6.8%	92.6%7.8%	98.0%2.6%	98.2%2.4%
Face-Name association recall accuracy[1]	14.4%21.6%	18.8%19.6%	26.9%28.2%	31.3%26.3%
Pattern Comparison (number of correct responses in 20 secs)	10.32.6	9.83.4	9.51.8	10.02.3
Trail-Making A (secs)	40.122.6	36.913.4	41.120.4	36.216.9
Trail-Making B (secs)	123.1108.1	89.651.6	106.352.0	93.134.9
SOPT Pattern (number of unique responses out of 16)	11.61.8	11.51.8	11.71.6	11.71.1
SOPT Word (number of unique responses out of 16)	14.21.7	13.91.6	14.11.1	14.41.4
MSEQ overall memory self-efficacy strength	60.916.0	58.712.0	62.723.6	60.319.4
EMQ number of memory errors	11.18.8	6.38.4	9.86.7	2.92.6
Word Memory Accuracy	N/A	69.5%10.0%	N/A	77.6%11.3%
Source Memory Accurcy	N/A	48.6%15.8%	N/A	54.3%15.4%

Note: SOPT = Self-Ordered Pointing Test; MSEQ = Memory Self-Efficacy Questionnaire (maximum score = 100); EMQ = Everyday Memory Questionnaire; DRM = Deese-Roediger-McDermott false memory test.

[1] Instructions were changed (from accepting "I don't know" as a response to requiring a guess) during the course of testing to try to avoid ceiling effects. The new instructions were used for 3 of the Integrated Sentences participants and all members of the Strategy Choice group. As can be seen, this instruction raised overall accuracy (though it was still very low), but did not influence transfer effects.