

Memory Updating Practice Across 100 Days in the COGITO Study

Yee Lee Shing

Max Planck Institute for Human Development and Humboldt-Universität zu Berlin

Florian Schmiedek

Max Planck Institute for Human Development and German Institute for International Educational Research (DIPF), Frankfurt am Main, Germany

Martin Lövdén

Max Planck Institute for Human Development and Karolinska Institutet and Stockholm University

Ulman Lindenberger

Max Planck Institute for Human Development

We investigated working memory updating performance in younger and older adults before, during, and after 100-day practice. Performance to presentation time (PT) relation was fitted to a negatively accelerated logistic function. Relative to younger adults, older adults showed lower asymptotes at pretest and posttest, and shallower slopes at pretest. Older adults practicing the task with fast PT gained less than older adults practicing the task with slow PT, probably reflecting the persistent use of a selective strategy throughout the 100-day practice period in the fast PT group. These results have implications for designing and evaluating age-comparative working memory training programs.

Keywords: aging, working memory, training, strategy use

Working memory (WM) commonly refers to a system or cognitive resource for the simultaneous storage and processing of information (Baddeley, 2003; Cowan, 1995; Just & Carpenter, 1992; Shah & Miyake, 1996). Numerous studies have shown that WM performance declines with advancing age (e.g., Park et al., 2002). Hence, WM training programs have gained popularity in aging research, with the goal of offsetting or delaying age-related cognitive deficits (see reviews in Hertzog, Kramer, Wilson, & Lindenberger, 2009; Lövdén, Bäckman, Lindenberger, Schaefer, & Schmiedek, 2010; Lustig, Shah, Seidler, & Reuter-Lorenz, 2009; Noack, Lövdén, Schmiedek, & Lindenberger, 2009).

Promising results from WM training programs with adult samples have recently been published by several research groups (e.g., Dahlin, Stigsdotter Neely, Larsson, Bäckman, & Nyberg, 2008; Jaeggi, Buschkuhl, Jonides, & Perrig, 2008; Li et al., 2008; Schmiedek, Lövdén, & Lindenberger, 2010). Cognitive training is

often successful at the group level for both younger and older adults, but individual differences in the amount of gain are large (Bissig & Lustig, 2007; Nyberg et al., 2003). So far, this phenomenon has primarily been observed in episodic memory training. In particular, older adults differ markedly in how much they profit from cognitive training, with the benefits found to be the smallest for individuals with lower initial cognitive status in some studies (magnification account; e.g., Bissig & Lustig, 2007; Singer, Lindenberger, & Baltes, 2003), while in other studies the opposite was found (compensation account; e.g., Carlson et al., 2008). The conditions under which magnification or compensation may occur are not well understood but may be related to types and forms of training. Understanding the reasons for these individual differences may help in the design of effective training programs that are tailored to individuals' cognitive status.

Here we examine practice of WM processes in younger and older adults with data from the COGITO study (Schmiedek, Bauer, Lövdén, Brose, & Lindenberger, 2010; Schmiedek, Lövdén, et al., 2010). The central aim of the COGITO study was to investigate day-to-day variability and plasticity of cognitive functioning in a broad multivariate way. To this end, a facet structure of cognitive tasks, cross-classifying perceptual speed, episodic memory, and WM with the content domains of verbal, numerical, and figural-spatial task material, was selected. Participants practiced 12 tasks (one WM, one episodic memory, and two perceptual speed tasks for each content domain) across 100 daily sessions, for an average of 1.0–1.5 hr per session. In this report, we focus on a task that targets a central WM process: updating. The updating function refers to the continuous modification of the content of WM according to incoming information (Miyake et al., 2000). Updating processing requires flexibility in information processing and a progressive shift of attention, that is, discarding no-longer-relevant

This article was published Online First October 10, 2011.

Yee Lee Shing, Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany, and Department of Psychology, Humboldt-Universität zu Berlin, Germany; Florian Schmiedek, Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany, and German Institute for International Educational Research (DIPF) Frankfurt am Main, Germany; Martin Lövdén, Max Planck Institute for Human Development, Berlin, Germany, and Aging Research Center, Karolinska Institutet and Stockholm University, Sweden; Ulman Lindenberger, Max Planck Institute for Human Development.

The first author (Shing) expresses gratitude to the fellowship support from the Alexander von Humboldt Foundation.

Correspondence concerning this article should be addressed to Yee Lee Shing, Center for Life Span Psychology, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. E-mail: yshing@mpib-berlin.mpg.de

information while new information is registered. Processes involved in updating predict WM capacity (Smith & Jonides, 1999) and complex span performance (Schmiedek, Oberauer, Wilhelm, Süß, & Wittman, 2007), as well as intelligence (Friedman et al., 2006). Relative to younger adults, older adults show impaired updating ability (Hartman, Dumas, & Nielsen, 2001), rendering this function a suitable candidate for intervention (see also Dahlin, Nyberg, Bäckman, & Neely, 2008). But in comparison to other cognitive functions, such as episodic memory and task switching, memory updating did not receive as much attention in the intervention literature, despite promising results from two studies that documented general positive effects of updating practice in older adults (Dahlin et al., 2008; Li et al., 2008). Therefore, to build on the emerging literature, we focused on the updating task of our broad task battery, particularly taking a microgenetic approach to provide more fine-grained insight into the training of memory updating in younger and older adults. In this vein, we want to point out that the choice of 100 days of practice should not be prescribed as the required dosage for training, as previous updating studies with shorter training phase (e.g., 5 weeks in Dahlin et al., 2008) demonstrated positive training results in older adults (but see Jaeggi et al., 2008, for effect of training dosage on transfer). Rather, the extensive training phase of the COGITO study was to allow for more fine-grained analyses of intraindividual variability, analogous to measurement-burst designs.

Based on both theoretical and empirical grounds, an important feature that training paradigms need to incorporate is adapting task difficulty to individuals' ability level (cf. Vygotsky, 1978). According to Lövdén and colleagues (Lövdén et al., 2010), the driving force of plastic alterations is a prolonged mismatch between functional capacity and experienced demands. Hence, demands need to be within the range of functional capacity to efficiently trigger plasticity, and the mismatch should be located between the extremes of being either too easy or too difficult.

In the COGITO study, careful consideration was given in presenting tasks at the appropriate difficulty level. This was attempted by adjusting the presentation time (PT) of the updating task based on time-accuracy function (TAF), which was assessed for each participant at pretest sessions, so that accuracy levels were above chance levels but still low enough to allow for a great amount of improvement without a ceiling being reached. For the updating task, participants were assigned to different PT practice conditions (i.e., 750 ms, 1500 ms, or 3000 ms) according to their TAF at baseline (see details in Method section). Individualized PTs were kept constant across the 100 sessions of the study, reflecting the investigators' interest in observing day-to-day fluctuations that are uncontaminated by variations in task difficulty. During the posttest sessions, participants were again assessed by TAF.

TAFs provide a joint description of speed and accuracy at the individual level that can be interpreted in general mechanistic terms. Accuracy is modeled as a negatively accelerated logistic function of PT. The function assumes (a) a minimum amount of PT is required to initiate processing, (b) beyond this minimum amount, PT is translated into carrying out the core cognitive processes necessary for achieving the task, and (c) after a certain maximum amount of PT, there is little to be gained by having more PT available (Kliegl, 1995; Kliegl, Mayr, & Krampe, 1994). In this sense, the parameter estimates derived from TAF separate process-

ing from asymptotic aspects of WM performance (cf. Norman & Bobrow, 1975).

So far, WM training studies have not distinguished training effects on processing and asymptotic components systematically. Some studies relied on accuracy measures, response time, or both, but without considering the relations between the two measures (e.g., Li et al., 2008). This comes with interpretational ambiguity, as differences between studies may reflect substantive variations in the construct of interest, in the scales or dimensions used, or both. The use of different measures renders attempts to integrate findings from the accuracy and time domains difficult, particularly because the relation between these two domains varies by task complexity and age (Kliegl et al., 1994; Verhaeghen, 2000). In contrast, TAF provides a joint time-accuracy platform for the investigation of age differences (cf. Cerella, 1990).

TAF is a nonlinear function. In this paper, TAFs were modeled using nonlinear mixed (NLME) models (Pinheiro & Bates, 2000). This analysis strategy carries the advantage that fixed effects (i.e., sample averages) and random effects (i.e., individual deviations from sample averages) representing TAF parameters are estimated simultaneously. Modeling TAF using NLME models is methodologically innovative because previous applications of TAF were generally limited to mean differences (except for Oberauer & Kliegl, 2006). By using NLME models, the parameters derived from the TAF can be modeled simultaneously as a function of quasi-experimental manipulations (i.e., age) and experimental manipulations (i.e., PT) at the sample level (i.e., fixed effects of age and PT), and as a function of interindividual differences (i.e., random effects).

The primary goals of this study were to examine both mean age and individual differences in effects of practice on WM updating in younger and older adults. We expected younger and older adults to differ in the asymptotic and processing components of memory updating. Based on positive training outcomes documented by previous studies, both age groups were expected to show gains in memory updating components after 100-day practice. However, we did not expect age differences in memory updating to be completely overcome. With more exploratory analyses, we also examined individual differences in practice gains as a function of the practice condition (based on PT) and individual differences in strategy use. The close examination of strategy use in task processing during training is a central, but often overlooked, component in evaluating the design of a training study. In addition, it provides useful clues on individual and age-related differences in task-relevant mechanisms (cf. Hertzog, Cooper, & Fisk, 1996). In this study, we identified and examined a selective strategy that was often used by the participants across the 100-day phase. We expected that the adaptive use of the selective strategy would relate to greater practice gains, as it allowed individuals to keep task difficulty in the medium (i.e., most performance-enhancing) range.

Method

Participants

Participants were recruited through newspaper advertisements, word-of-mouth recommendation, flyers distributed in university buildings, community organizations, and local stores. The advertisements targeted people who were interested in practicing cog-

nitive tasks for 4–6 days a week for a period of about 6 months. Financial remuneration was mentioned, but no details were given regarding amounts in the advertisements. There were no explicit exclusion criteria for the study. However, several steps were taken to maximize the chance that we included participants who were willing and capable of remaining with the study throughout. First, interested people were given information about the study in a telephone interview and inquiries were made as to whether the requirements for participation in the study, in particular, the time investment, could be met. Potential candidates were then invited to join a 1-hr “warm-up” group session whereby the general aims of the study were explained and detailed information on incentives was given.

During the practice phase of the study, 101 younger (51.5% women; age: $M = 25.6$ years, 20–31 years) and 103 older adults (49.5% women; age: $M = 71.3$ years, 65–80 years) completed an average of 101 practice sessions (younger adults: $M = 100.8$, $SD = 2.6$, range = 87–109; older adults: $M = 101.0$, $SD = 2.7$, range = 90–106). Older adults were administered the MMSE at pretest and posttest and had at least a total score of 25 on one of the two occasions. Furthermore, both younger and older samples were relatively representative regarding general cognitive functioning, as indicated by comparisons of Digit Symbol performance with data from a population-based study and a meta-analysis (see details in Schmiedek, Lövdén, et al., 2010). More descriptive characteristics about the sample can be found in Table 1. Further information about the study is reported in previous publications (Schmiedek, Bauer, et al., 2010; Schmiedek, Lövdén, et al., 2010).¹

Attrition rate for participants who had entered the longitudinal practice phase was low (i.e., 15 out of 219 participants; for details on rates and reasons of dropout in the different study phases, see Schmiedek, Bauer, et al., 2010).

Procedure

Before and after the longitudinal phase, participants completed pretest and posttest in 10 sessions that consisted of 2–2.5 hours of comprehensive cognitive test batteries and self-report questionnaires. On average, the time elapsing between pretest and posttest was 197 days for the younger and 188 days for the older adult groups, respectively.

During the longitudinal practice phase, participants scheduled daily sessions (1.0–1.5 hours) on an individual basis for up to 6 days a week (including Saturdays). Participants worked on the tasks individually in rooms with three to six workstations. At the end of each session, participants received feedback on their own performance on all tasks, including average accuracies and reaction times. They were also able to receive printouts of these results to take home. Discussing results with staff members was not an official part of the procedure. But if participants had questions, the staff members were available and open to talk about them.

Memory Updating Task

In this adapted version of numerical memory updating (Salt-house, Babcock, & Shaw, 1991), four single digits (ranging from 0 to 9) were presented simultaneously for 4000 ms in four horizontally arranged cells. After an ISI of 500 ms, a sequence of eight

Table 1
Descriptive Characteristics of Sample

Measure	Younger Adults Mean (SD)	Older Adults Mean (SD)
Age	25.6 (2.7)	71.3 (4.1)
Gender (Female %)	52%	50%
Occupational status		
Attending school/university	76%	7%
Working	13%	0%
Jobless	11%	1%
Retired	0%	92%
Years of education		
High school	12.5 (1.3)	10.8 (1.8)
Post high school	3.6 (2.8)	2.9 (2.8)
Cognitive status		
WAIS digit-symbol	60.3 (9.5)	43.6 (9.0)
Accuracy Spot-a-Word test	0.7 (0.1)	0.8 (0.1)
Health status		
Health complaints	0.6 (0.4)	0.6 (0.4)
Number of illness	1 (1.1)	3.6 (2.4)
Self-rated sleep quality	1.9 (0.7)	1.9 (0.6)

Note. Scale for health complaints ranged from 0 (none), 1 (little), 2 (some), 3 (considerable), to 4 (strong), averaged across 33 kinds of complaints (e.g., headache, lethargy, memory loss). Number of illnesses referred to the sum of “yes” response across 31 kinds of illnesses (e.g., neurological disorder, cancer). Sleep quality ranged from 0 (very bad), 1 (quite bad), 2 (quite good), to 3 (very good).

updating operations was presented in a second row of four cells (directly below the first row of four cells where the digits appeared). These updating operations were additions and subtractions within a range of -8 to $+8$. They had to be applied to the digits memorized from the corresponding cells above, and the updated results had to be memorized. Possible PTs of updating operation were 750, 1500, 3000, and 6000 ms (in blocks). ISI was 250 ms. At the end of each trial, the four end results had to be typed into the four cells.

The same test version was used for both pretest and posttest, except that at pretest, participants were given a practice block with 6 trials on each PT. For the actual test, there were 12 trials for each PT, divided into two blocks. The sequence of block was fixed to be 6000 ms, 3000 ms, 1500 ms, 750 ms, 750 ms, 1500 ms, 3000 ms, and, finally, 6000 ms. By using the same test version, we made sure that the tests at pretest and posttest did not differ in difficulty or any other subtle characteristic. However, possible retest effect should be kept in mind when interpreting the results.

For the 100-day practice, eight blocks were included in each daily session. In between blocks, participants were allowed to make breaks at their own pace and start the next block by themselves. The median time spent on the updating task was about 3.5 min for the younger and 4.5 min for the older adults (from when the first block began until the results of the last block were entered). Participants were assigned into practice conditions of either having 750, 1500, or 3000 ms for updating operation based

¹ The COGITO study also consisted of a no-training control group, with 44 younger (age 21–29 years) and 39 older adults (age 65–81 years). Detailed information about the control group, particularly in regard to training gain and transfer effect, can be found in Schmiedek, Lövdén, et al. (2010).

on their pretest performance. Pretest performance was estimated individually by fitting TAF to each person's data. Specifically, participants who were assigned into the fastest practice condition (750 ms) performed at pretest above 32.5% for the 750-ms condition or above 55% in the 1500-ms condition at pretest. Participants assigned into the 1500-ms practice condition performed between 32.5% and 55% at pretest in the 1500-ms condition, or above 55% in the 3000-ms condition. The rest of the sample was assigned to the 3000-ms practice condition. In total, 63 younger adults were assigned to the 750-ms condition, 33 to the 1500-ms condition, and 5 to the 3000-ms condition. For older adults, 30 were assigned to the 750-ms condition, 52 to the 1500-ms condition, and 21 to the 3000-ms condition.

Results

**TAF at Baseline and Posttest:
Age and Practice Effects**

As mentioned earlier in the article, NLME was used as a data analysis tool for fitting TAF. We used SAS PROC NL MIXED. The relation between time (pt) and accuracy (p) can be modeled as a negatively accelerated function (e.g., Lohman, 1989; Wickelgren, 1977). In particular, we modeled the TAFs (simultaneously for each age group at pretest and posttest) with the following general logistic function:

$$p = d + (a - d)/(1 + \exp(- (c + b*pt))) \quad (1)$$

where *d* is a parameter for chance performance; *a* is the asymptotic maximum accuracy; *b*, scaled in terms of time, represents the steepness of the function and thus the efficiency of processing over time (i.e., the rate with which the asymptote is approached); and

$-c/b$ represents the inflection point at which changes in PT lead to maximum changes in accuracy. For the present analysis, *d* is fixed to 0.10 because there were 10 potential answers (0 through 9). In contrast to previous studies, we modeled the TAF as logistic instead of negative exponential functions. The two functions are relatively similar but the logistic function provided better fits to the current data.

The estimated TAFs of each age group at pre- and posttests are plotted against PTs in Figure 1. The estimates of each TAF parameter (fixed effects) are summarized in Table 2. To examine the effects of age and practice, models in which relevant parameters were constrained to be equal were estimated (see model comparisons in Table 3). All models were estimated with a log-likelihood fitting function (-2LL). The constrained models were nested within the freely estimated models, such that the -2LL difference is chi-square distributed with degrees of freedom equal to the difference in the number of parameters. Hence, chi-square tests were used to compare nested models. The alpha level was set to $p = .01$, corresponding to a critical chi-square value of 6.63 for 1 degree of freedom.

Overall, younger and older adults showed significantly different asymptotes and slopes at pretest, such that younger adults reached higher levels of asymptotic performance and were more responsive to the increase of PT. Both age groups improved significantly in asymptotes and slopes after 100 days of practicing the task. At posttest, older adults still reached significantly lower asymptotic performance, but their mean slope was no longer different from younger adults. Examining the difference curves in estimated accuracy between the two age groups helps in identifying practice-induced shifts in age group differences. At pretest, the group of older adults performed below the level of younger adults across short and long PTs. At posttest, age group differences were con-

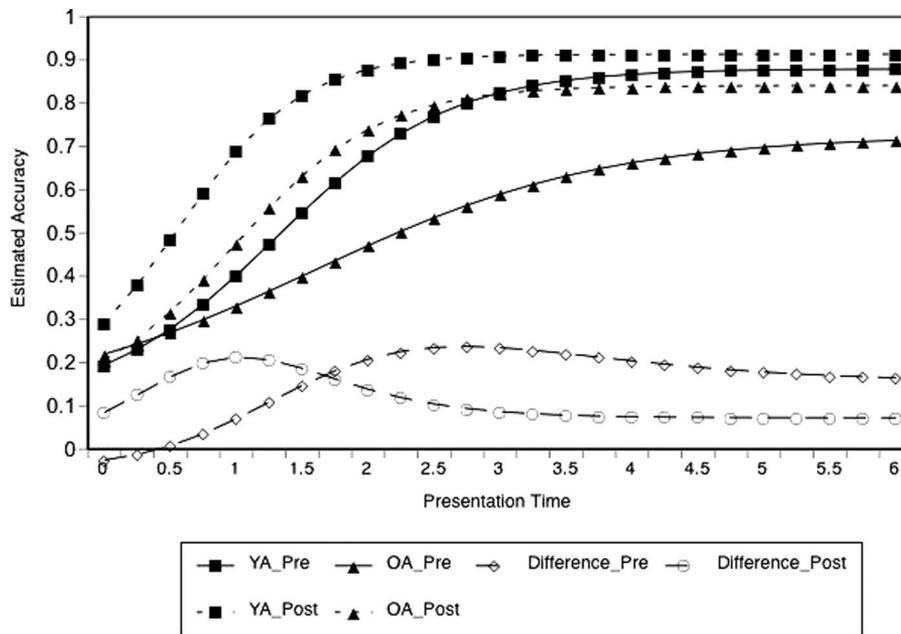


Figure 1. Estimated time-accuracy function and difference curve of younger and older adults before and after 100-day practice.

fined to very short PTs. Thus, practice narrowed the performance gap between younger and older adults without closing it.

Evaluating Gain by Practice Condition

Given that participants were assigned to three different PT conditions for the 100-day practice phase, we also estimated the TAFs separately by condition. Due to the small number of participants assigned to the 3000-ms condition, stable estimation of TAF parameters could not be achieved and the group was not included in the analyses reported here. For illustration, the TAF plots of participants assigned to the 750-ms (fast group) and 1500-ms (slow group) practice conditions are presented in Figure 2, separately for younger and older adults.

As observed in Figure 2, participants in the different practice groups differed greatly in pretest performance, reflecting the non-random nature of the group assignment. The largest group differences were observed in the middle range of PT, that is, between 1 and 3 s of PT. In either age group, differences in asymptotes between the two practice groups appeared no longer reliable at posttest. In the group of older adults, differences in slopes vanished as well. For younger adults, posttest slope differences as a function of practice group seemed to confine to very short PT (i.e., less than 1 s).

To formally capture these patterns, we further examined the performance of different groups (defined by the crossing of age and practice group) separately for each PT using a latent difference score model (LDSM) in a multiple group setting (see Figure 3; cf. McArdle & Nesselroade, 1994). In the model, the improvement in performance between pretest and posttest is captured by a latent difference variable whose mean, variance, and covariance are estimated separately for the four groups (i.e., younger adult/fast, younger adult/slow, older adult/fast, older adult/slow groups). The corresponding parameter estimates are presented in Table 4. Between-groups differences in parameters were again tested with nested model comparisons.

Regarding age differences within practice groups, there was no difference in performance between younger and older adults for their pretest performance in the condition they were assigned to. This finding suggests that the assignment of individuals into practice groups worked well in ensuring that people within each group performed similarly at baseline. However, younger and older adults differed in how much they gained from the practice. Specifically, younger adults gained more than the older counterparts, both within the fast (.29 vs. .06, $\Delta\chi^2 = 34.33, df = 1$) and slow (.41 vs. .28, $\Delta\chi^2 = 18.78, df = 1$) practice groups.

Table 2
Fixed Effect Estimates of TAF Parameters for Younger and Older Adults at Pretest and Posttest

	Asymptote	Slope	Inflection Point
Pretest			
Younger adults	0.88 (.01)	1.52 (.08)	1.30
Older adults	.73 (.02)	.91 (.07)	1.58
Posttest			
Younger adults	.91 (.01)	2.15 (.18)	.55
Older adults	.84 (.01)	1.82 (.09)	.98

Table 3
Results of Model Comparison to Test TAF Parameters for Age and Practice Effects

Model	-2LL
Free estimation	-3538
Age Effect	
Pretest	
asymptote YA = asymptote OA	-3501*
slope YA = slope OA	-3500*
Posttest	
asymptote YA = asymptote OA	-3521*
slope YA = slope OA	-3534
Practice Effect	
YA	
Asymptote baseline = Asymptote posttest	-3529*
Slope baseline = Slope posttest	-3526*
OA	
Asymptote baseline = Asymptote posttest	-3504*
Slope baseline = Slope posttest	-3471*

* $p < .01$.

When examining differences in initial performance and practice gains within age groups as a function of practice condition, we confirmed that participants in the fast practice group performed better than participants in the slow practice group at pretest for both 750 ms and 1500 ms, again reflecting the nonrandom assignment of participants to practice groups. For 1500-ms PT, participants in the slow practice group gained significantly more than those in the fast practice group, leading to a reduction of performance differences with practice. This was the case for both younger adults (.41 vs. .18, $\Delta\chi^2 = 47.92, df = 1$) and older adults (.28 vs. .07, $\Delta\chi^2 = 40.99, df = 1$). This finding is not surprising given that those trained on the slow (i.e., 1500 ms) condition should gain more in the 1500-ms condition. However, a different pattern emerged in the 750-ms condition. Both younger and older adults of the fast practice group (i.e., 750 ms) showed practice gains that did not differ reliably from the practice gains of their age peers in the slow practice group ($\Delta\chi^2 < 6.63, ns$). Taken together, participants assigned to the fast practice group did not gain as much from practice participants assigned to the slow practice group. For younger adults, this might be related to their relatively high performance already at pretest, which may have limited the room for further improvement. However, this explanation is unlikely to hold for older adults in the fast practice group, who showed smaller gains in performance than older adults in the slow practice group despite sufficient room for improvement.²

² We also estimated the LDSM including the no-training control group of younger and older adults. When comparing the difference score estimate, indeed older adults in the fast practice condition did not show significantly higher gain than the older adults control group (for PT = 750 ms, $M_diff = .04$; for PT = 1500 ms, $M_diff = .03$; compare estimates in Table 4). However, younger adults in the fast and slow practice conditions, as well as older adults in the slow practice condition, showed significantly higher gain than their age peers in the control group. This additional analysis reinforced our findings that gain in the older adult group was mainly driven by participants in the slow practice condition.

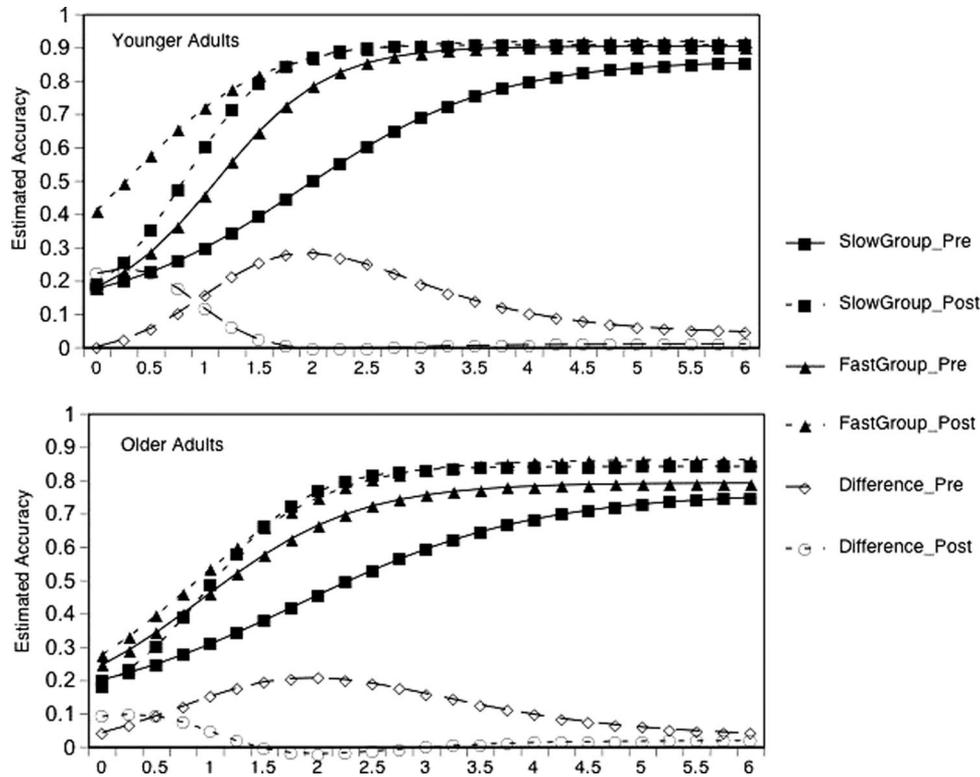


Figure 2. Estimated time-accuracy function and difference curve of younger and older adults segregated by fast and slow practice condition, before and after 100-day practice.

Differential Use of Selective Strategy by Practice Group

To examine the reason why older adults in the fast group did not gain much from practice, we turned to the daily data and took a closer look at how participants approached the task during the practice period. The updating task was set up with four cells (presented in a horizontal array), in which arithmetic operations appeared sequentially in a random manner and digits within the four cells had to be processed and updated. This set up makes it

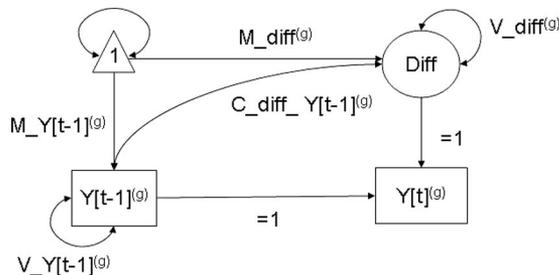


Figure 3. Latent difference score model implemented in a multiple-group setting (as indicated with g) to capture baseline performance, gain in performance from baseline to posttest, and the relation between the two. M_{diff} = estimated mean difference in performance (from pretest to posttest); $M_Y[t-1]$ = estimated pretest performance; $C_{diff_Y[t-1]}$ = correlation between pretest performance and mean difference in performance.

possible for participants to only process a subset of the four cells to render the task more manageable (e.g., ignore the digit and updating operations of one or several of the cells). Thus, we analyzed individuals' performance for each of the four updating cells to explore the possibility that participants in the two practice groups differed in strategy use. We rank ordered individuals' performance for the four cells from best to worst, separately for each day across the practice phase. If a participant used the selective strategy, the performance of the ignored cell should be close to chance and/or clearly lag behind the performance of the other cells in which the participant paid attention to. The four panels in Figure 4 are performance of four individuals who showed different patterns of change. Participant A, an older adult in the slow group, demonstrated equal improvement in performance in all four cells, suggesting that he or she paid attention to updating all cells in a similar manner. Participant B, a younger adult in the slow group, initially focused only on the updating operation of two cells. But roughly at the midpoint of practice (e.g., after approximately 50 sessions), he or she showed sharp improvements in the third cell, while the performance for the cell with the lowest performance did not change much throughout. We called these individuals *decreasing selectors*. Participant C, an older adult in the fast group, showed little gain in performance across all four cells. Finally, Participant D, a younger adult in the fast group, persistently showed good performance for two cells but low performance for the other two cells, thereby showing a selective preference for two of the four cells. We called these individuals *fixed selectors*.

Table 4
Estimates From Latent Difference Score Model

	M_diff	M_Y[t-1]	C_diff_ Y[t-1]
Presentation time = 750 ms			
YA-slow	.21	.27	-.81
YA-fast	.29	.37	-.57
OA-slow	.12	.27	-.56
OA-fast	.06	.40	-.64
Presentation time = 1500 ms			
YA-slow	.41	.39	-.39
YA-fast	.18	.66	-.39
OA-slow	.28	.40	-.48
OA-fast	.07	.60	-.43

Note. Numbers in italic refer to conditions in which participants were trained on. M_diff = estimated mean difference in performance (from pretest to posttest); M_Y[t-1] = estimated pretest performance; C_diff_ Y[t-1] = correlation between pretest performance and mean difference in performance.

To formally capture the pattern of each group, we modeled the relation between time (day) and accuracy of best- to worst-performing cells (rank-ordered each day by performance) separately as a negatively accelerated logistic function, similar to the one used for the TAF. The estimated trajectories of each group are illustrated in Figure 5. As can be seen, older adults in the fast practice group clearly lagged behind in their third and worst performing cells, with the worst cell being at the chance level throughout the practice phase. This result suggests that these older adults might have ignored the updating operations of one of the four cells through the practice phase. Younger adults in the fast group (and, to a lesser degree, older adults in the slow group) also showed a gap in performance between their worst cell and the other three cells, but the gap was not as large compared to the one of the older adults in the fast group.

An alternative way to rank order the cells is to take the average of each cell from left to right position, as they were presented in a horizontal array, across all days. In contrast to the previous analysis, in which the identity of cells could change from day to day, this alternative approach retains the identity of the cells across sessions for a given person. Analyzing the data in these two different ways addresses, to some degree, the question of whether the use of the selective strategy was intentional, as the second approach would suggest a coherent and possibly intentional strategy, while the first approach could also be seen as a nonstrategic byproduct of neglect due to lack of resources (i.e., it was simply too taxing to update four cells). The two ways of analyzing the data yielded very similar results. In both analyses, older adults in the fast group clearly lagged behind in their third and worst performing cells, with the worst cell at chance level of performance throughout the practice phase. Similarly, younger adults in the fast group also showed a gap in performance in their worst cell. Also, accuracy followed a left-to-right pattern, in the sense that cells further to the right of the array were associated with lower levels of performance.³ In sum, the selection of which cell to pay attention to for performing the updating operations was largely coherent.

Discussion

In this study, we reported a set of analyses in which the relation between PT and memory updating performance was modeled as a

negatively accelerated logistic function, before and after 100-day practice, in younger and older adults. The two age groups showed significantly different asymptotes and slopes at pretest, indicating that both the processing and asymptotic components of memory updating were compromised in older adults relative to younger adults. Thus, older adults were less capable of making use of additional time available for processing and showed reduced upper limits of performance when given maximum amounts of time for processing. After 100 days of practicing the task, asymptotes and slopes had improved in both age groups. Age group differences in asymptotic performance persisted at posttest, pointing to an age-related decline in memory updating that was not overcome by 100 days of practice.

Based on pretest performance, participants were assigned to practice conditions that differed in PT. When inspecting practice gains as a function of PT condition, we found that participants assigned to the fast condition (750 ms) gained less from practice than participants assigned to the slow condition (1500 ms). For younger adults in the fast practice condition, room for improve-

³ To further explore support for an adaptation-based interpretation of condition differences in performance gains without the confound of initial difference in WM, we identified a subgroup of older adults from each practice condition who showed closest levels of performance at pretest for the updating task at PT of 750 ms and 1500 ms. In particular, the 12 highest performing older adults in the slow condition and the 11 lowest performing older adults in the fast condition were selected. At pretest, the two selected subgroups did, in fact, not differ reliably from each other for the 750-ms condition ($p = .6$), while for the 1500-ms condition, the fast subgroup (.53) performed slightly better than the slow subgroup (.46), $F(1, 21) = 4.11$, $p = .06$. Examining specific PT condition at posttest, for the 1500-ms condition, the 12 selected older adults in the slow group (.68) tended to perform better than the 11 selected age peers in the fast group (.59), $F(1, 21) = 3.26$, $p = .09$, despite the initial higher performance of the fast subgroup at pretest for this condition. No other significant or marginally significant difference was found for the other PT conditions. When examining the use of selective strategy in these subgroups of older adults, it was found that the patterns resembled the full sample. Namely, participants in the slow subgroup tended to leave out two cells in the beginning but gradually included one more cell, whereas participants in the fast subgroup tended to leave out two cells throughout the 100 days of practice. In sum, differential use of selective strategy was found in these subgroups of older adults, despite the matching of pretest performance.

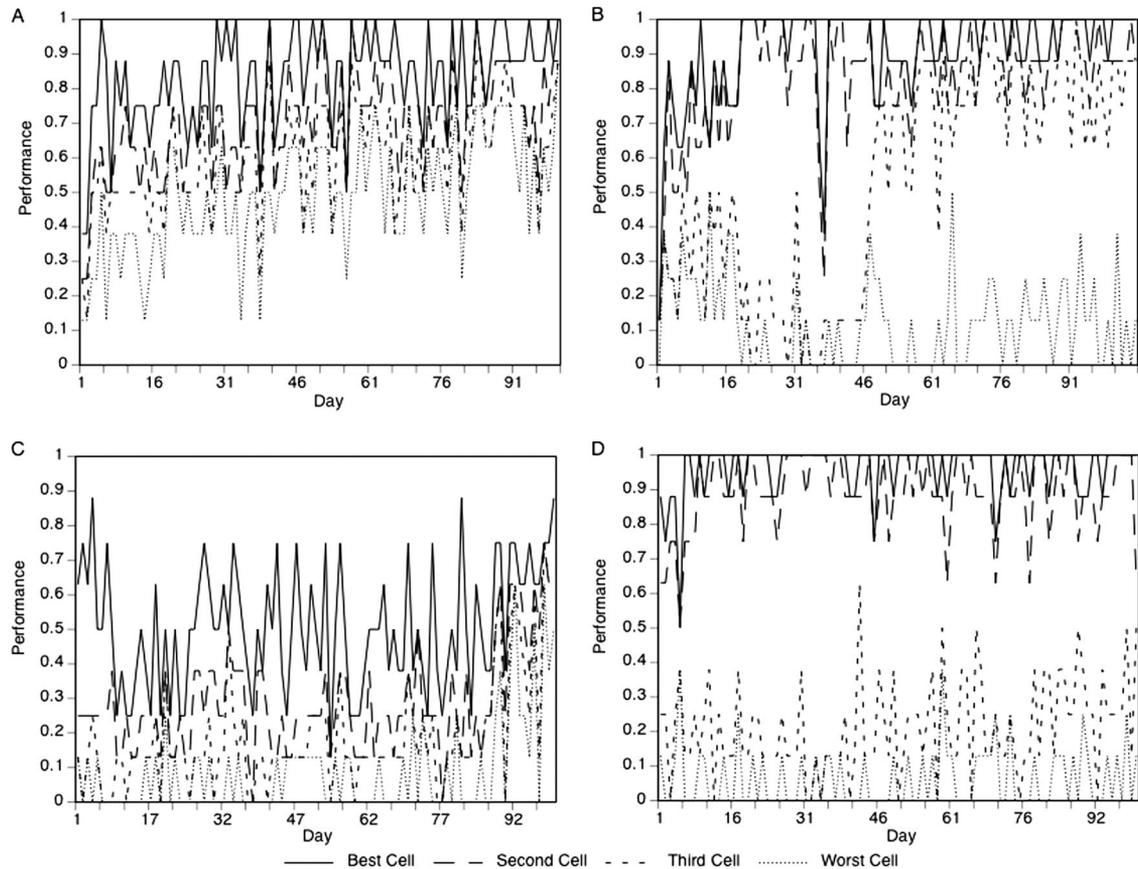


Figure 4. Examples of typical pattern of performance change from best to worst cells from four individuals. Panel A represents an individual who showed improvements in performance for all four cells; Panel B represents a *decreasing selector* who focused on two out of four cells but progressed to focus on three out of four cells after half of the practice phase; Panel C represents an individual who showed overall little improvements for all four cells; and Panel D represents a *fixed selector* who selectively focused on only two out of the four cells throughout the practice phase.

ment may have been limited due to high levels of pretest performance. This explanation does not hold for older adults in the fast-practice condition. In their case, the persistent use of a selective strategy throughout the 100-day phase may have resulted in less frequent and demanding drilling of updating operations in the course of practice. In contrast, older adults in the slow PT condition also tended to focus on only two cells of the task in the initial period of practice but gradually included more cells when working on the task. The interpretation that differences in practice regime contributed to differences in performance gains beyond initial differences in performance was supported by exploratory subgroup analyses based on older adults who showed similar performance at pretest (see Footnote 3). Again, older adults assigned to the slow practice group attained slightly better updating performance at posttest than those assigned to the fast practice group, and also showed less fixated use of selective strategy. These results have implications for our previous report that demonstrated training gain and transfer effects in a latent WM factor for the practice group (Schmiedek, Lövdén, et al., 2010). Given that participants were assigned into different PT conditions for all three WM tasks, it is important to further examine to what extent gain and transfer

effects in the older adults were driven by participants of certain PT conditions.

We note similarities between the present approach and the microgenetic method advocated by Siegler and others to study the adaptive nature of children's variability in strategy use (see review in Siegler, 1996). The microgenetic method involves obtaining frequent samples of children's behavior as their cognition is undergoing change. In this article, we demonstrate that the consideration of strategy use in task processing is also important in evaluating practice paradigms and gains, especially for understanding the cognitive mechanisms that are intended for practice. While the use of a selective strategy in the updating task may seem less adaptive at first glance, we should keep in mind that individuals engage in selection of strategies (either automatically or consciously) for a task depending on the performance goals of the individual, the affordances of the task context, and the individual's capabilities. The selective strategy may indeed be adaptive by holding difficulty of the task at a maximally manageable level for the individual. Moreover, this strategy can be adapted over time so that task difficulty is kept at the maximally manageable level and therefore enhances the demand characteristics of the practice re-

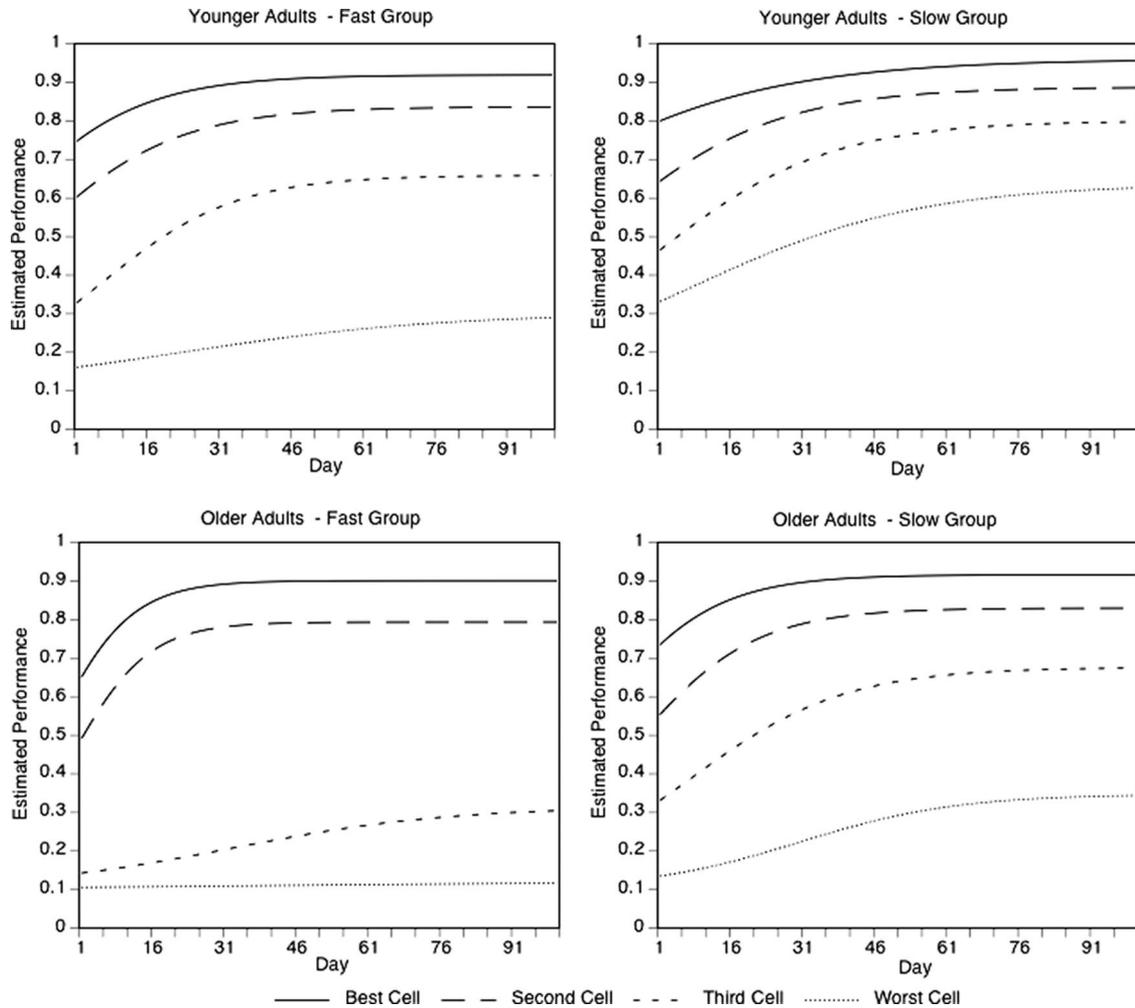


Figure 5. Pattern of performance change from best to worst cells for younger and older adults in the fast and slow conditions.

game (not too difficult and not too easy; cf. Lövdén et al., 2010). From this point of view, the decreasing selectors (see Figure 4, panel B) were using the most adaptive strategy. The fixed selectors (see Figure 4, panel D), on the other hand, were initially using the same strategy but failed to eventually include more cells in their processing. The group analysis showed that most of the older adults in the fast condition behaved like fixed selectors.

There are several possible reasons for the extensive use of the selective strategy by older adults assigned to the fast condition. In the skill acquisition literature, there is a substantial amount of evidence indicating that the strategies used by older adults in a variety of learning situations are less optimal than those used by younger adults (Hertzog et al., 1996; Rogers & Gilbert, 1997; Touron & Hertzog, 2004a, 2004b). In a task in which participants verify whether a centrally presented target noun pair matches one of a set of pairs contained in a lookup table at the top of the screen, Touron and Hertzog (2004a, 2004b) reported that older adults were more reluctant to shift from a scanning strategy to the retrieval of noun pairs during the course of skill learning. This age difference remained even when

older adults had sufficient knowledge of the noun-pair associations to support the retrieval-based strategy, arguing against an age-based learning deficit as the reason for older adults' reluctance to shift strategy. Rather, Touron and Hertzog (2004b) showed that age differences in strategy selection were reliably associated with older adults' lack of confidence in using the retrieval strategy, pointing to the importance of metacognitive factors. It is possible that, while the older adults assigned to the fast condition had initial higher WM capacity, the fast presentation of stimuli created a situation that taxed the upper limit of their functional range, both objectively and subjectively. In this group, accuracies for the two cells with the lowest and second-to-lowest scores were consistently below 30%. These low levels of performance might have affected their subjective perception of how they were performing in the task, which influenced their subsequent choice of strategies. We agree that this study did not provide conclusive evidence about the degree to which strategy choices were deliberate or incidental. However, the consistent rank-ordering of cells across sessions by array position renders it likely that the selective strategy observed in this study was intentional and goal-directed, and not just an automatic consequence of scarce re-

sources. The present paradigm seems well suited to explore this question in future studies.

This study has several limitations that should be considered. First, the assignment of individuals to practice conditions was nonrandom, confounding the interpretation of group differences in strategy use and performance gains. The subgroup analyses served to attenuate this shortcoming. Second, although we demonstrated that older adults in the fast condition were behaving like fixed selectors, we did not attempt to relate patterns of strategy use to learning within individuals. Third, we analyzed data from one task of a 12-task practice battery. Exposure to the other tasks might have influenced performance on the memory updating task, and the exact mechanism of practice is difficult to tease apart. Rather, the unique aspect of our current analyses is to complement previous promising results of updating training (e.g., Dahlin et al., 2008; Li et al., 2008) by taking a more fine-grained approach to examine individual differences in strategy use. The use of such selective strategy has not been documented in the WM training literature and we think that our article serves to underscore the importance of considering strategy use factors. Our results are also in line with the call for implementing alternative research designs, such as measurement-burst designs, for gaining more fine-grained understanding of mechanisms of cognitive enrichments (see Hertzog et al., 2009).

To situate our study within the broader context of cognitive enrichment in the public domain, the external validity of our results and scalability of the training paradigm remains to be demonstrated. Despite huge commitment required from the participants, the attrition rate was remarkably low in the COGITO study. Although this is a desirable characteristic of interventions, there are several qualifications to be considered. The first issue is that our sample may be more positively selected than the general population. Cognitively, we demonstrated that our younger and older samples were quite representative, as indicated by comparisons of Digit Symbol performance with data from a population-based study and a meta-analysis (Schmiedek, Lövdén, et al., 2010). However, our training group—in particular, the older adults—might indeed be more motivated than those who chose not to participate to remain cognitively intact. Another consideration is the strong financial incentives the participants received. For many participants, the increase in social contact and gains in self-efficacy were also important factors for keeping them committed to the program (see Schmiedek, Bauer, et al., 2010). These are factors that need to be taken into consideration when designing cognitive intervention to ensure programs are sustainable and accessible to the larger public.

Furthermore, as the older adults of our main study were mostly elderly with relatively intact cognitive functioning, the clinical implication of the training is not immediate. However, participants in our main study were recruited again for a 2-year follow up, which included measurements on all cognitive tasks and more detailed health assessments. In addition, about 90% of the older adults and 50% of the younger adults also started to participate in the Socio-Economic Panel (SOEP), a representative longitudinal household survey in Germany conducted by the German Institute for Economic Research. With several longitudinal follow-ups on both the practice and control groups, we anticipate to have more appropriate data to evaluate the clinical utility of our training

procedure, including the possibility of postponing decline and pathological aging.

In conclusion, the results of this study show that 100 days of practice improved older adults' memory updating performance up to the level shown by younger adults at baseline. Older adults assigned to the slow PT condition brought about most of these gains. Older adults in the fast condition might have been confronted by an overly difficult task and apparently settled on a strategy that was less effective in promoting performance gains. These results have general implications for the design and evaluation of cognitive interventions in old age. Although task difficulty was adjusted individually based on baseline performance, our study revealed the shortcoming of not adapting task difficulty across the practice phase dynamically. Dynamic adjustment of task difficulty should be the goal of practice and training, as it serves to promote learning by continuously calibrating task difficulty at a challenging, yet manageable, level (e.g., Brehmer, Li, Müller, von Oertzen, & Lindenberger, 2007; Klingberg et al., 2005). Our results also suggest that microgenetically assessing and monitoring older adults' approach to the cognitive task they practice is important. Furthermore, a closer examination of metacognitive factors, which include motivational, self-related, and volitional dimensions, may enhance the efficiency of training programs and provide a more complete and dynamic picture of cognitive change.

References

- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Review Neuroscience*, *4*, 829–839. doi:10.1038/nrn1201
- Bissig, D., & Lustig, C. (2007). Who benefits from memory training? *Psychological Science*, *18*, 720–726. doi:10.1111/j.1467-9280.2007.01966.x
- Brehmer, Y., Li, S.-C., Müller, V., von Oertzen, T., & Lindenberger, U. (2007). Memory plasticity across the lifespan: Uncovering children's latent potential. *Developmental Psychology*, *43*, 465–478. doi:10.1037/0012-1649.43.2.465
- Carlson, M. C., Saczynski, J. S., Rebok, G. W., Seeman, T., Glass, T. A., McGill, S., . . . Fried, L. P. (2008). Exploring the effects of an everyday activity program on executive function and memory in older adults: Experience Corps. *The Gerontologist*, *48*, 793–801. doi:10.1093/geront/48.6.793
- Cerella, J. (1990). Aging and information processing rates in the elderly. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 201–221). New York, NY: Academic Press.
- Cowan, N. (1995). *Attention and memory: An integrated framework*. New York, NY: Oxford University Press.
- Dahlin, E., Nyberg, L., Bäckman, L., & Neely, A. S. (2008). Plasticity of executive functioning in young and older adults: Immediate training gains, transfer, and long-term maintenance. *Psychology and Aging*, *23*, 720–730. doi:10.1037/a0014296
- Dahlin, E., Stigsdotter Neely, A., Larsson, A., Bäckman, L., & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science*, *320*, 1510–1512. doi:10.1126/science.1155466
- Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*, 172–179. doi:10.1111/j.1467-9280.2006.01681.x
- Hartman, M., Dumas, J., & Nielsen, C. (2001). Age differences in updating working memory: Evidence from the delayed-matching-to-sample test. *Aging, Neuropsychology, and Cognition*, *8*, 14–35. doi:10.1076/anec.8.1.14.847
- Hertzog, C., Cooper, B. P., & Fisk, A. D. (1996). Aging and individual

- differences in the development of skilled memory search performance. *Psychology and Aging*, *11*, 497–520. doi:10.1037/0882-7974.11.3.497
- Hertzog, C., Kramer, A. F., Wilson, R. S., & Lindenberger, U. (2009). Enrichment effects on adult cognitive development: Can the functional capacity of older adults be preserved and enhanced? *Psychological Science in the Public Interest*, *9*, 1–65.
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 6829–6833. doi:10.1073/pnas.0801268105
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149. doi:10.1037/0033-295X.99.1.122
- Kliegl, R. (1995). From presentation time to processing time: A psychophysics approach to episodic memory. In F. E. Weinert & W. Schneider (Eds.), *Memory performance and competencies* (pp. 89–110). Mahwah, NJ: Erlbaum.
- Kliegl, R., Mayr, U., & Krampe, R. T. (1994). Time-accuracy functions for determining process and person differences: An application to cognitive aging. *Cognitive Psychology*, *26*, 134–164. doi:10.1006/cogp.1994.1005
- Klingberg, T., Fernell, E., Olesen, P., Johnson, M., Gustafsson, P., Dahlström, K., . . . Westerberg, H. (2005). Computerized training of working memory in children with ADHD: A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*, 177–186. doi:10.1097/00004583-200502000-00010
- Li, S.-C., Schmiedek, F., Huxhold, O., Röcke, C., Smith, J., & Lindenberger, U. (2008). Working memory plasticity in old age: Practice gain, transfer, and maintenance. *Psychology and Aging*, *23*, 731–742. doi:10.1037/a0014343
- Lohman, D. F. (1989). Estimating individual differences in information processing using speed-accuracy models. In R. Kanfer, P. L. Ackerman & R. Cudeck (Eds.), *Abilities, motivation, and methodology* (pp. 119–163). Hillsdale, NJ: Erlbaum.
- Lövdén, M., Bäckman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychological Bulletin*, *136*, 659–676. doi:10.1037/a0020080
- Lustig, C., Shah, P., Seidler, R., & Reuter-Lorenz, P. A. (2009). Aging, training, and the brain: A review and future directions. *Neuropsychology Review*, *19*, 504–522. doi:10.1007/s11065-009-9119-9
- McArdle, J. J., & Nesselroade, J. R. (1994). Using multivariate data to structure developmental change. In S. H. Cohen & H. W. Reese (Eds.), *Life-span developmental psychology: Methodological contributions* (pp. 223–266). Hillsdale, NJ: Erlbaum.
- Miyake, A., Freidman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100. doi:10.1006/cogp.1999.0734
- Noack, H., Lövdén, M., Schmiedek, F., & Lindenberger, U. (2009). Cognitive plasticity in adulthood and old age: Gauging the generality of cognitive intervention effects. *Restorative Neurology and Neuroscience*, *27*, 435–453. doi:10.3233/RNN-2009-0496
- Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7*, 44–64. doi:10.1016/0010-0285(75)90004-3
- Nyberg, L., Sandblom, J., Jones, S., Stigsdotter Neely, A., Petersson, K. M., Ingvar, M., & Bäckman, L. (2003). Neural correlates of training-related memory improvement in adulthood and aging. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 13728–13733. doi:10.1073/pnas.1735487100
- Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory & Language*, *55*, 601–626.
- Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuo-spatial and verbal memory across the adult life span. *Psychology and Aging*, *17*, 299–320. doi:10.1037/0882-7974.17.2.299
- Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effect models in S and S-plus*. Berlin, Germany: Springer.
- Rogers, W. A., & Gilbert, D. K. (1997). Do performance strategies mediate age-related differences in associative learning? *Psychology and Aging*, *12*, 620–633. doi:10.1037/0882-7974.12.4.620
- Salthouse, T. A., Babcock, R. L., & Shaw, R. J. (1991). Effects of adult age on structural and operational capacities in working memory. *Psychology and Aging*, *6*, 118–127. doi:10.1037/0882-7974.6.1.118
- Schmiedek, F., Bauer, C., Lövdén, M., Brose, A., & Lindenberger, U. (2010). Cognitive enrichment in old age: Web-based training programs. *GeroPsych*, *23*, 59–67. doi:10.1024/1662-9647/a000013
- Schmiedek, F., Lövdén, M., & Lindenberger, U. (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: Findings from the COGITO study. *Frontiers in Aging Neuroscience*, *2*, 1–10. doi:10.3389/fnagi.2010.00027
- Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., & Wittman, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, *136*, 414–429. doi:10.1037/0096-3445.136.3.414
- Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, *125*, 4–27. doi:10.1037/0096-3445.125.1.4
- Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York, NY: Oxford University Press.
- Singer, T., Lindenberger, U., & Baltes, P. B. (2003). Plasticity of memory for new learning in very old age: A story of major loss? *Psychology and Aging*, *18*, 306–317. doi:10.1037/0882-7974.18.2.306
- Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, *283*, 1657–1661. doi:10.1126/science.283.5408.1657
- Touron, D. R., & Hertzog, C. (2004a). Distinguishing age differences in knowledge, strategy use, and confidence during strategic skill acquisition. *Psychology and Aging*, *19*, 452–466. doi:10.1037/0882-7974.19.3.452
- Touron, D. R., & Hertzog, C. (2004b). Strategy shift affordance and strategy choice in young and older adults. *Memory & Cognition*, *32*, 298–310. doi:10.3758/BF03196860
- Verhaeghen, P. (2000). The parallels in beauty's bow: Time-accuracy functions and their implications for cognitive aging theories. In T. J. Perfect & E. A. Maylor (Eds.), *Models of cognitive aging* (pp. 50–86). New York, NY: Oxford University Press.
- Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
- Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, *41*, 67–85.

Received January 4, 2011

Revision received July 18, 2011

Accepted July 22, 2011 ■