Preservation of memory-based automaticity in reading for older adults

By: Katherine A. Rawson and Dayna R. Touron

Abstract:

Concerning age-related effects on cognitive skill acquisition, the modal finding is that older adults do not benefit from practice to the same extent as younger adults in tasks that afford a shift from slower algorithmic processing to faster memory-based processing. In contrast, Rawson and Touron (2009) demonstrated a relatively rapid shift to memory-based processing in the context of a reading task. The current research extended beyond this initial study to provide more definitive evidence for relative preservation of memory-based automaticity in reading tasks for older adults. Younger and older adults read short stories containing unfamiliar noun phrases (e.g., skunk mud) followed by disambiguating information indicating the combination's meaning (either the normatively dominant meaning or an alternative subordinate meaning). Stories were repeated across practice blocks, and then the noun phrases were presented in novel sentence frames in a transfer task. Both age groups shifted from computation to retrieval after relatively few practice trials (as evidenced by convergence of reading times for dominant and subordinate items). Most important, both age groups showed strong evidence for memory-based processing of the noun phrases in the transfer task. In contrast, older adults showed minimal shifting to retrieval in an alphabet arithmetic task, indicating that the preservation of memory-based automaticity in reading was task-specific. Discussion focuses on important implications for theories of memory-based automaticity in general and for specific theoretical accounts of age effects on memory-based automaticity, as well as fruitful directions for future research.

Article:

Research on cognitive skill acquisition has consistently shown that older adults do not benefit from practice to the same extent as younger adults. These age differences are particularly pronounced for tasks in which practice affords a shift from slower algorithmic processing to faster memory-based processing, with older adults consistently slower or less likely to shift to memory-based processing during cognitive skill acquisition (e.g., Hoyer, Cerella, & Onyper, 2003; Jenkins & Hoyer, 2000; Rogers, Hertzog, & Fisk, 2000; Touron & Hertzog, 2004a, 2004b; Touron, Hoyer, & Cerella, 2004; Touron, Swaim, & Hertzog, 2007). However, Rawson and Touron (2009) reported an intriguing empirical exception to this modal finding, in which older adults demonstrated a relatively rapid shift to memory-based processing in a reading task. The current research extended beyond this initial study to provide more definitive evidence for relative preservation of memory-based automaticity in reading tasks for older adults.

Below, we first briefly describe memory-based theories of automaticity. Next, we summarize research on age effects on memory-based processing in cognitive skill acquisition. Last, we discuss evidence reported by Rawson and Touron (2009) and how the current research provides a significant extension to definitively establish robust memory-based automaticity in reading tasks for older adults. To foreshadow, the Discussion focuses heavily on the theoretical implications of these outcomes, both for theories of memory-based automaticity in general and for specific theoretical accounts of age effects on memory-based automaticity.

MEMORY-BASED THEORIES OF AUTOMATICITY

Early theories of automaticity involved what Logan (1988) referred to as property-list accounts, in which automaticity is defined in terms of properties of task performance (e.g., task performance might be defined as "automatic" if it is fast, effortless, and autonomous; e.g., Hasher & Zacks, 1979; Schneider, Dumais, & Shiffrin, 1984). This approach to defining automaticity is also common in theories of reading skill, in which "skilled reading" is often defined in terms of properties of performance (e.g., Walczyk, 2000), and in theories of specific component processes within the comprehension system (e.g., word recognition, T. L. Brown, Gore, & Carr, 2002; syntactic parsing, Hahne & Friederici, 1999; inferencing, McKoon & Ratcliff, 1992). However, as Logan and Klapp (1991) stated, "approaches that define automaticity in terms of manifest properties, such as speed, effortlessness, and autonomy (property-list approaches), are stipulative or descriptive but not predictive" (p. 179), because they do not explain the underlying processes that give rise to these properties. Rawson (2004) also noted that property-list accounts seldom agree on the set of necessary and sufficient properties for defining reading as "skilled" or "automatic." Given these limitations of property-list accounts of automaticity, contemporary theories of skill acquisition instead define automaticity in terms of the dynamic underlying processes that give rise to the performance properties of interest.

For present purposes, we focus on memory-based theories of automaticity (Logan, 1988; Palmeri, 1997; Rickard, 1997). These theories assume that many tasks afford two possible routes for interpreting stimuli, an algorithmic route and a memory-based route. Algorithms involve an item-general sequence of processing steps to interpret any stimulus of a given type (e.g., multiplication rules can be applied to solve any multiplication problem). In contrast, the memory-based route involves retrieving interpretations of stimuli previously stored in long-term

memory (e.g., directly retrieving "$12 \times 12 = 144$"). The core assumption of memory-based theories is that speed-ups with practice reflect shifting away from slower algorithmic processes to greater reliance on retrieval of prior interpretations. For example, instance theory (Logan, 1988) assumes that each time a stimulus is encountered, the algorithmic and memory-based routes race to produce an interpretation, and stimulus interpretation is based on the output of whichever process finishes first. The likelihood that at least one instance will be retrieved quickly increases as the number of instances increases; thus, retrieval is increasingly likely to produce a response before the algorithm.

A common method for detecting the role of memory-based automaticity in speed-ups with practice (and the one adopted here) involves comparing response times at the beginning versus end of practice for stimulus sets that differ in their algorithmic complexity. For example, Logan and Klapp (1991) presented learners with alphabet arithmetic problems (e.g., F + 2 = H?). At the outset of practice, individuals solve problems algorithmically by counting up the alphabet from the initial letter to the terminal letter. Problems with larger versus smaller addends (e.g., F + 4 = J vs. F + 2 = H) take longer to solve via counting. Later in practice, individuals are increasingly likely to directly retrieve solutions stored on prior trials, and retrieval speed is assumed not to differ by addend size. For example, in Logan and Klapp (1991), response times increased by 486 ms per increase in addend at the beginning of practice versus by only 45 ms at the end of practice. More generally across tasks used in automaticity research, the degree of convergence in response times at the end of practice for items that initially differ in algorithm complexity indicates the degree that performance involves memory-based processing.

The contribution of memory-based processing to speed-ups with practice has been shown in a wide range of skill acquisition tasks, including alphabet arithmetic (e.g., Klapp, Boches, Trabert, & Logan, 1991; Logan & Klapp, 1991; Wilkins & Rawson, 2010, 2013), pound arithmetic (described below; e.g., Rickard, 1997; Touron et al., 2004), multiplication (e.g., Rickard, Lau, & Pashler, 2008), visual search (e.g., Fisk, Hertzog, Lee, Rogers, & Anderson-Garlach, 1994), noun-pair look-up (e.g., Rogers et al., 2000; Touron & Hertzog, 2004a), lexical decision (e.g., Grant & Logan, 1993; Logan, 1988), and numerosity judgments (e.g., Jenkins & Hoyer, 2000; Lassaline & Logan, 1993; Palmeri, 1997). Most relevant for present purposes, recent research has also established the contribution of memory-based processing to speed-ups with practice in syntactic and semantic processes during reading tasks (Rawson, 2004, 2010; Rawson & Middleton, 2009; Rawson & Touron, 2009).

AGE DIFFERENCES IN MEMORY-BASED AUTOMATICITY

Although research has demonstrated memory-based automaticity in a wide range of skill acquisition tasks, these effects are not equally robust for younger and older adults. Across many different tasks, older adults are slower or less likely to shift to memory-based processing with practice (e.g., Hoyer et al., 2003; Jenkins & Hoyer, 2000; Rogers et al., 2000; Touron & Hertzog, 2004a, 2004b; Touron et al., 2004; Touron et al., 2007). For example, Touron et al. (2004) presented older and younger adults with pound arithmetic verification problems (of the abstract form A # B = C, in which "#" refers to the following sequence of operations: B − A + 1 + B = C; e.g., "12 # 43 = 75" is true, because 43 – 12 + 1 + 43 = 75). Ten problems were each presented for 75 practice trials, and participants were prompted to report how they verified the preceding

problem (by computing the answer or retrieving the answer). Despite the small number of problems and the large number of practice trials, reports of retrieval use were still lower for older adults than for younger adults at the end of practice. Other studies have also shown persistent age differences in memory-based automaticity even after extensive practice with the same items (e.g., Hertzog, Touron, & Hines, 2007; Jenkins & Hoyer, 2000; Lamson & Rogers, 2008; Touron, 2006). For example, Touron and Hertzog (2004a) found persistent age differences in retrieval use after 60 item repetitions, even in a condition for which all items were pretrained.

In contrast to the overwhelming evidence for persistent age-related deficits in memory-based automaticity, Rawson and Touron (2009, Experiment 1) reported an apparent exception. They repeatedly presented younger and older adults with short stories containing unfamiliar noun phrases (e.g., *bee caterpillar*) that afforded a normatively preferred semantic interpretation (referred to as the *dominant* meaning, e.g., *a caterpillar that looks like a bee*) and an interpretation that was generated infrequently in prior norming studies (referred to as the *subordinate* meaning, e.g., *a caterpillar that buzzes like a bee*). In each story, the ambiguous noun phrase was followed by a sentence indicating the combination's intended meaning (referred to as the *disambiguating region,* which contained either the dominant or subordinate meaning; see sample materials in Table 1).

Table 1
Sample Materials Used in the Reading Practice and Reading Transfer Tasks

Sample materials

Narratives from reading practice task:

Chris had very eccentric tastes when it came to decorating his apartment. What an unusual collection of pieces he had assembled! On one wall, he had a huge felt painting of dogs playing poker. Hanging on another wall over a blood-red leather couch were two lavishly decorated Samurai swords. And in front of the shelf housing his lava lamp collection was one of his favorite pieces. The tiger chair he bought on eBay had cost him a pretty penny. He liked to brag about the chair [*that was entirely covered in tiger print/that had been used by famous tiger trainers*]. Chris thought he had designed quite a bachelor pad, although his friends snickered about the fact that he hadn't had a real date in several months.

Kelly was becoming increasingly sorry that she had agreed to take the neighbor boy along on the family camping trip. They hadn't been able to send their son Seth to summer camp this year, so she told him he could bring a friend camping to make it up to him. If only she had known what she was getting herself into! Just today, the boys had pushed Seth's little sister into the lake, ripped a hole in the tent, and almost set the dog on fire. She was at her wit's end when she realized that the boys had discovered the skunk mud down by the lake. They had been playing in the mud that [*was as stinky as a skunk/had a dead skunk in it*]. With no running water at the campsite, she didn't know how to even begin cleaning up.

Sentences from reading transfer task:

Sensible: Candice was pleased with the tiger chair she had added to her living room as an accent piece because [*it was entirely covered in tiger print/ it had been used by famous tiger trainers*].

Nonsensical: Jill's dog had rolled around in some skunk mud, and as she bathed him she thought that the mud [*that was as stinky as a skunk/that had a dead skunk in it*] would never wash out of his shaggy limousine.

Note. For illustrative purposes here, the novel noun phrase in each text is underlined and the disambiguating region is bracketed and italicized. In the experiment, the disambiguating region contained either the dominant meaning (the first phrase in the bracket) or the subordinate meaning (the second phrase in the bracket).

On the initial encounter of an unfamiliar phrase, no prior interpretations are stored in memory and, thus, semantic processes must generate an interpretation (for theories of the processes involved in interpreting novel noun phrases, e.g., see Estes & Glucksberg, 2000; Gagné & Shoben, 1997; Gerrig & Bortfeld, 1999; Middleton, Rawson, & Wisniewski, 2011; Murphy, 2002; Wisniewski, 1997; Wisniewski & Middleton, 2002). These computational processes will presumably generate the dominant meaning, which by definition is the one most frequently generated in norming studies. If the subsequent disambiguating region contains the dominant meaning, no reanalysis will be needed. If the disambiguating region contains the subordinate meaning (which by definition is a meaning infrequently generated in norming studies), the initially generated meaning will be inconsistent with the intended meaning and reanalysis will be required. Reanalysis in the disambiguating region elevates reading times, and thus, longer reading times in the subordinate versus dominant condition were expected at the outset of

practice. Consistent with this expectation, reading times during the first presentation trial were significantly longer in the subordinate versus dominant condition for both age groups (on average, 252 ms longer for younger adults and 422 ms longer for older adults).

The key prediction concerns reading times in the disambiguating region in later blocks of practice. Upon reencounter of the noun phrases in later presentations of the stories, the correct interpretations stored on previous trials can be retrieved in both conditions, avoiding the need for reanalysis in the subordinate condition. Consistent with this account, reading times in the subordinate versus dominant condition did not significantly differ during the last presentation trial for either age group (a 9-ms difference for younger adults and a 13-ms difference for older adults). Although older adults were somewhat slower to shift to memory-based processing (reading times in the subordinate and dominant conditions converged statistically by Trial 3 for younger adults but not until Trial 7 for older adults),[1] the finding that older adults completely shifted to memory-based processing after only seven trials stands in stark contrast to prior research involving nonreading tasks, in which older adults typically demonstrate only partial shifting at best even after extended amounts of practice (upward of 60–100 trials in many cases).

To our knowledge, these outcomes reported by Rawson and Touron (2009) represent the only demonstration of preserved memory-based automaticity in older adults. Other research has demonstrated that both younger and older adults show fluency gains when rereading lengthy passages, although they differ somewhat in the level of processing showing facilitation (e.g., Stine-Morrow, Gagne, Morrow, & DeWall, 2004). However, these and other rereading methodologies (for a review, see Collins & Levy, 2008) do not pin down the involvement of memory-based processing per se, given that fluency gains may also reflect improvement in the efficiency of algorithmic processes (e.g., Wilkins & Rawson, 2010). Given that Rawson and Touron (2009) report the only direct evidence for relatively preserved memory-based automaticity in older adults, one goal of the current research was to replicate these initial findings (for recent emphasis on the importance of replicating novel findings, see Pashler & Harris, 2012; Roediger, 2012; Schmidt, 2009).

More important, the current research included two novel extensions to provide more definitive evidence of robust memory-based automaticity for older adults during reading. Of key interest, the current research evaluated the extent to which older adults show preserved memory-based automaticity when practiced items are presented in novel contexts. Stories containing noun phrases disambiguated with either their dominant or subordinate meaning were presented repeatedly during the practice phase (as in Rawson & Touron, 2009), followed later by a transfer phase in which the noun phrases were presented in new sentence contexts in a different task (see examples in Table 1). It is important to note that the new sentence context either disambiguated with the same meaning from practice or with the other meaning. If interpretation of the noun phrase in the transfer sentence is based on retrieval of previously stored interpretations, reading times in the subsequent disambiguating region will be a function of match with the practiced meaning. If the practiced and transfer meanings are the same, no reanalysis will be needed in the

---

[1] Rawson and Touron (2009) found no age difference in the rate of shift in Experiment 2 (reading times for both age groups statistically converged by Trial 2). However, Experiment 2 involved a pretraining phase in which participants learned the meanings for some of the combinations prior to practice; thus, these outcomes should be interpreted with some caution concerning age similarities in spontaneous shifting to memory-based automaticity.

disambiguating region. If the practiced and transfer meanings are not the same, reanalysis will be needed and, thus, reading times will be slower. In a previous study using this methodology (Rawson & Middleton, 2009), younger adults demonstrated this pattern indicating transfer of memory-based processing, but the extent to which older adults will show similar levels of transfer is unknown. If memory-based processing is less robust for older adults, retrieval of prior interpretations of the noun phrases may be minimally involved during transfer. If so, interpretation will revert back to the algorithm, which is most likely to generate the dominant meaning. In this case, only a main effect of the meaning in transfer would be predicted, with faster reading times for transfer items that disambiguate with the dominant meaning versus the subordinate meaning.

Additionally, between each block of practice trials in the reading task, the secondary task involving repeated presentation of alphabet arithmetic problems (as described earlier). We included this task to demonstrate that our older adult sample shows the more typical pattern of deficits in a nonreading task, which is important for establishing that any preservation of memory-based automaticity observed in the reading tasks is specific to reading rather than reflecting task-general preservation.

In sum, the current research provides several important contributions to the literature on age effects on memory-based processing in cognitive skill acquisition. As a starting point, we provide a replication of Rawson and Touron's (2009) initial findings, which is important given that their single outcome stands in stark contrast to the wealth of prior research demonstrating marked age effects on memory-based automaticity. As increasingly emphasized in the field, replication of novel findings is important in its own right (e.g., Pashler & Harris, 2012; Roediger, 2012; Schmidt, 2009), particularly for unexpected or counterintuitive outcomes. Additionally, this replication involves a stronger test of older adults' relative preservation of memory-based automaticity during reading by examining shifting to memory-based processing under somewhat more difficult practice conditions due to increases in set size and lag here versus in Rawson and Touron (2009), two factors that have been shown to retard the shift to memory-based processing (Rickard et al., 2008; Wilkins & Rawson, 2013, 2014).

More important, the current work also provides two novel extensions beyond Rawson and Touron (2009) that are key for establishing patterns of age deficits and sparing of memory-based automaticity. First, we more clearly establish preservation of memory-based automaticity in reading by examining age equivalence when items are transferred to novel contexts, which must be established before any strong claims can be made about the generality of this mechanism in reading for older adults. The possibility remains that memory-based interpretation for older adults' may have been contextually bound to the repeated contexts in which the novel phrases occurred (cf. Levy & Burns, 1990). Second, we provide the first demonstration of both age deficits and sparing of memory-based automaticity as a function of task in the same sample of older adults. As discussed further below, these outcomes have important implications for both general theories of automaticity as well as the dominant theories of age effects on memory-based automaticity (associative deficit hypothesis and retrieval aversion hypothesis).

METHOD

Participants and Design

Participants included in the analyses reported below were 55 younger adults (69% female female) and 59 older adults (68% female).[2] Younger adults were students enrolled in psychology courses at Kent State University who participated for course credit. Older adults were community residents from the Greensboro, North Carolina, area who received a $30 honorarium for their participation. Demographic data and performance on the baseline cognitive measures are reported in Table 2. Age groups differed in education level, digit–symbol performance and vocabulary. However, these variables were not systematically related to the response time outcomes of interest in the reading or alphabet arithmetic (AA) tasks (almost all associations were small and not significant), and thus we do not discuss them further.

Table 2
Demographic Information and Performance on Baseline Cognitive Measures

Measures	Younger adults			Older adults		
	M	*SD*	range	*M*	*SD*	range
Demographic variable						
Age (years)	19.4	1.8	17–26	66.7	4.3	60–75
Education*	12.9	1.3	11–16	16.3	0.5	12–21
Health rating	1.8	0.8	1–4	1.6	0.5	1–3
Baseline cognitive measure						
Digit–symbol (in 90 s)*	62.1	8.5	43–81	40.1	8.1	24–62
Digit–symbol recall (out of 9)*	7.1	2.2	1–9	6.0	2.2	0–9
Vocabulary (out of 40)*	27.1	3.7	20–37	35.1	3.4	27–40

Note. Education = self-reported number of years of formal education completed. Health rating = self-reported rating of overall health as compared to a perfect state of health, from 1 (*very good*) to 5 (*very poor*). Digit–symbol, digit–symbol recall, and vocabulary are reported as the number of questions answered correctly; see text for descriptions of these tasks.
* Difference between younger and older adults significant at $p < .01$.

In the reading practice task, practice meaning (subordinate vs. dominant) and practice block (1–8) were within-participant manipulations. In the reading transfer task, practice meaning (subordinate vs. dominant) and transfer meaning (subordinate vs. dominant) were within-participant manipulations. In the alphabet arithmetic task, addend size (2 vs. 4) and practice block (1–8) were within-participant manipulations.

Materials

Reading tasks

Target items for the reading practice task and the reading transfer task included 24 novel noun-noun combinations (hereafter referred to as *combinations* for brevity). Each combination had a *dominant* meaning and a *subordinate* meaning. As in the vast majority of prior research on conceptual combination, the dominant and subordinate meanings were defined empirically based on response frequencies from participants in norming studies who were asked to generate a meaning for each novel combination. The dominant meaning is defined as the meaning most

---

[2] An additional eight older adults and six younger adults were tested but excluded from analyses because they did not complete all experimental tasks, had performance below 70% on the comprehension questions during the main reading task, and/or had more than 40 "TOO FAST" warnings during the reading task (indicating noncompliance with task instructions to read carefully).

frequently generated by participants, and the subordinate meaning is defined as a plausible alternative meaning that was infrequently generated by participants. The items used here were previously developed via norming studies with both younger and older adults (Rawson & Middleton, 2009; Rawson & Touron, 2009).

In the reading practice task, each combination was embedded in a short narrative text (developed by Rawson & Middleton, 2009) that contained two critical sentences (see Table 1). The first critical sentence introduced the novel combination but did not explicitly state the intended meaning of the combination. The second critical sentence followed immediately after and contained a phrase explicitly stating the intended meaning of the combination; this phrase is referred to as the *disambiguating region*. Two versions of the disambiguating region were written for each sentence, one stating the dominant meaning and one stating the subordinate meaning. The disambiguating regions in the dominant and subordinate versions of each sentence were matched as closely as possible for number of letters (Ms = 25.7 and 24.8), number of syllables (Ms = 7.9 and 8.0), mean log word frequency across all words and across only content words (Ms = 4.5 and 4.2; Ms = 2.7 and 2.3; G. D. A. Brown, 1984). For each participant, 12 of the narratives were presented with the dominant version of the disambiguating sentence and 12 included the subordinate version, with assignment of narrative to condition counterbalanced across participants. Materials for the reading practice task also included eight *yes/no* comprehension questions for each experimental narrative, with a different question presented after each trial for a given narrative. Each question tapped understanding of nontarget material within the narrative. Materials also included five filler narratives used for instructions and warm-up at the beginning of practice, with a *yes/no* comprehension question for each.

Materials for the reading transfer task included 24 transfer sentences (developed by Rawson & Middleton, 2009). Each sentence contained a target combination that had been presented during the practice task. Transfer sentences were different than the sentences containing the combinations during the practice task (see Table 1). For each transfer sentence, the combination was mentioned in the early part of the sentence and disambiguated in the later part of the sentence. Two versions of the disambiguating region were written for each transfer sentence, one stating the dominant meaning and one stating the subordinate meaning. It is important to note that in each transfer sentence, the disambiguating information was either consistent or inconsistent with the meaning of the combination that had been presented during the reading practice task. That is, half of the combinations disambiguated with their dominant meaning during practice were disambiguated with their dominant meaning during transfer, whereas the other half were disambiguated with their subordinate meaning during transfer. Likewise, half of the combinations that were disambiguated with their subordinate meaning during practice were disambiguated with their subordinate meaning during transfer and the other half were disambiguated with their dominant meaning. Assignment of combinations to the two transfer conditions (dominant vs. subordinate meaning during transfer) was counterbalanced across participants.

Materials also included 48 filler sentences. Because the nominal task during transfer involved making sensibility judgments, half of the transfer sentences and half of the filler sentences were sensible, and the remaining sentences were nonsensical (see Table 1). Nonsensical sentences were created by replacing the final word in the sentence with a word that was semantically

inappropriate in that context (only the sentence-final word was altered to avoid influencing reading times in the preceding disambiguating region of the target sentences).

Alphabet arithmetic task

Materials included eight target problems, four with addend 2 and four with addend 4. At each addend size, two of the problems were true (e.g., $F + 2 = H$) and two were false (e.g., $E + 4 = J$). Six additional filler problems from a nonoverlapping portion of the alphabet were used for warm-up trials during instructions.

Procedure

Participants first answered basic demographic questions. They were then told that the goal of the experiment was to examine how practice in a reading task influences how efficiently people can learn from text. They were told they would read each text one segment at a time and were instructed to read each story carefully and to answer as many comprehension questions correctly as possible. As part of the instructions, participants were presented with two filler narratives to familiarize them with the moving window procedure. Before beginning the main task, they were told they would be given opportunities for short breaks at various points in the experiment (to encourage participants not to stop in the middle of a task).

The reading practice task began with presentation of the three remaining filler narratives as warm-up trials, followed by the first block of presentation trials for the target narratives in random order. Each narrative was presented using a moving window procedure that presented one region of each sentence at a time. Regions were primarily based on natural phrase boundaries and contained from 1–7 words (e.g., "As the day wore on, / Sarah started to wonder / if a black cat / had crossed her path"). The first region was presented in the upper left of the screen. When the participant pressed the spacebar, each alphanumeric character was replaced with dashes (spaces and punctuation were not replaced) and the next region was presented to the right of the first, and so on for each subsequent region. Participants could not move backward to reread previously viewed regions. To discourage skimming, if the mean reading time for 10 consecutive regions was less than 200 ms, the narrative was temporarily removed from the screen and the following warning was displayed for 4 s: "TOO FAST!! Please read each text carefully. The story will continue in a moment." Narrative presentation then resumed at the same point as before the warning. The computer recorded the position and number of warnings for each participant.

At the end of a narrative, the computer displayed a *yes/no* comprehension question. After incorrect responses, participants saw "ERROR! Please try to read each text carefully" for 1.5 s. A different comprehension question was presented after each trial for each narrative. Comprehension questions were included to support the cover task instructions that the experiment was about text learning and to encourage participants to read each narrative carefully. Performance on these questions was relatively high for younger and older adults ($M = 91\%$ and 93%, respectively) and will not be discussed further, given that it is not of theoretical interest for present purposes.

After the first block of trials in the reading practice task, participants were given brief instructions about the AA task, including presentation of the six filler problems to familiarize them with the procedure. Participants were encouraged to answer each problem correctly but as quickly as possible. Participants then completed the first block of AA trials, which included four repetitions of each of the eight target problems (items were each presented once in random order within each of four miniblocks). Each trial began with a row of asterisks in the middle of the screen for 500 ms, which was then replaced with the AA problem. Participants responded by clicking on one of two buttons below the problem (labeled *TRUE* and *FALSE*). The problem and response buttons were then replaced by a button labeled *NEXT* that participants clicked to initiate the next trial.

At the end of the first AA block, participants were offered a short break if needed. They then completed the second block of trials for the target narratives in the reading practice task, followed by the second block of AA trials, and so on until eight alternating blocks of reading practice and AA trials were completed. Thus, the 24 items in the reading task were each presented eight times during practice (once in each block), and the eight AA items were each presented 32 times during practice (four times in each block). Participants were offered a short break after each block of AA trials; on average, younger and older adults took less than one minute for breaks ($M = 0.30$ and $M = 0.55$ min).

Participants then completed a vocabulary test (adapted from the Shipley Vocabulary test, Zachary, 1986). Participants had 4 min to answer 40 multiple-choice questions in which they were to select the correct synonym for each prompt word. Participants then completed a digit symbol task (adapted from Wechsler Adult Intelligence Scale Digit Symbol Subtest; Wechsler, 1981), in which the Digits 1–9 were paired with symbols at the top of the screen. On each trial, a symbol was shown and participants responded with the corresponding number as quickly as possible. Participants had 90 s to answer as many questions as possible. In the digit–symbol recall task administered immediately after, the nine symbols were presented alone and participants recalled which digit went with each symbol.

The vocabulary and digit symbol tasks were administered not only to collect standard baseline cognitive measures, but also to help mask the purpose of the transfer task that followed. In the transfer task, participants were given cover task instructions that the next task was a measure of reading comprehension ability and were not explicitly told about the connection between the stimuli included in the transfer task and the previous reading practice task. Participants were instructed that they would be presented sentences one at a time to read at their own pace, and that "After you have read a sentence, you will be asked to judge whether the sentence was sensible or insensible. You will need to read each sentence carefully in order to make correct judgments. Your goal is to make as many correct judgments as possible." The sensibility judgment task was used to reduce the nominal similarity of practice and transfer tasks.

Participants were presented with sentences one at a time. The sentence list included the 24 transfer sentences and 48 filler sentences. Each sentence was presented one phrase at a time via the moving window procedure. After each sentence, participants clicked on one of two buttons to indicate whether the sentence was sensible or not. If a participant made an incorrect judgment, the computer displayed the word *ERROR* on the screen for 1.5 s before beginning the next trial.

RESULTS

Reading Practice Task

For each participant, we computed mean reading time in the disambiguating region within each block of trials for each condition. For these and all other reading time measures reported below, less than 1% of reading times for individual trials were removed as outliers.[3] Figure 1 reports mean reading times in the disambiguating region. Results of an omnibus 2 (age group) × 2 (subordinate vs. dominant meaning) × 8 (practice block) mixed-factor analysis of variance (ANOVA) are reported in Table 3. The three-way interaction was significant, which we examine further below by focusing discussion on outcomes from paired comparisons that evaluate directional predictions motivated by the results of previous research. For these and other comparisons, Cohen's d values were computed using pooled standard deviations (Cortina & Nouri, 2000).

As in prior research, both age groups showed practice effects, with faster reading times during the eighth versus first block of practice in the subordinate condition—younger adult: $t(54) = 19.72$, $p < .001$, $d = 3.13$; older adult: $t(58) = 20.43$, $p < .001$, $d = 3.40$—and in the dominant condition—younger adult: $t(54) = 20.77$, $p < .001$, $d = 3.25$; older adult: $t(58) = 18.29$, $p < .001$, $d = 2.98$. Of greater interest, comparisons of reading times for subordinate and dominant items in each block of practice provide evidence that these practice effects reflected increasing involvement of memory-based processing in both age groups. As reported in Table 4, reading times in the first block of practice were significantly slower in the subordinate versus dominant condition for both age groups, as expected. To revisit, all combinations are unfamiliar during Block 1, and thus, the dominant meaning is assumed to be computed initially in both conditions. In the subsequent disambiguating region, reanalysis will be necessary in the subordinate condition, as reflected by the longer reading times. In contrast, reading times for subordinate and dominant items converged statistically by Block 4 for younger adults and by Block 6 for older adults, with minimal differences thereafter (see effect sizes for later blocks of practice, Table 4). Upon later encounters of the combinations, both younger and older adults presumably shifted away from computation to retrieval of appropriate interpretations and, thus, avoided the need for reanalysis in the subordinate condition.

---

[3] We used a two-stage approach to exclude outliers. In Stage I, any region reading time <100 ms or >10,000 ms in any block was excluded. In Stage II, we identified individual mean values >3 SDs above group mean. Of 4,140 individual mean values inspected, 27 were outliers by this criterion. In 12 of these cases, removal of one long response time in that condition (ranging from 5,250–8,367 ms, all in later blocks of practice) brought the participant's mean value within the range of the group. In total across both stages, < 1% of reading times were excluded. We used the same approach for the alphabet arithmetic task. In Stage I, response times <50 ms or >1,2000 ms were excluded; in Stage II, 11 individual mean values were outliers, and all 11 were brought into the group range with the removal of 1–2 additional response times (from 9,134–11,684 ms).
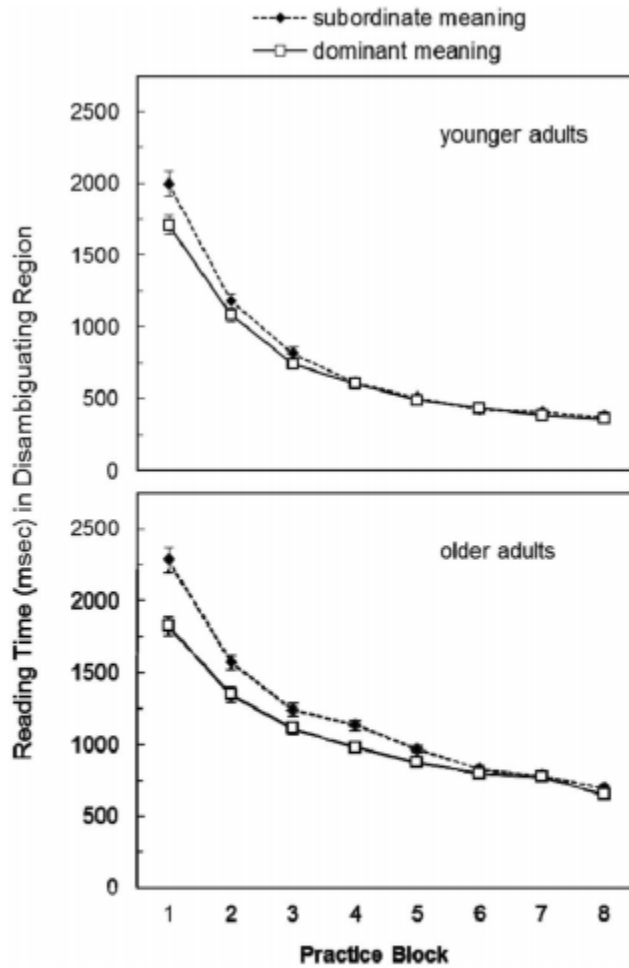
Figure 1. Mean reading time (in milliseconds) in the disambiguating region of target sentences in the reading practice task, for younger adults (top panel) and older adults (bottom panel), as a function of the disambiguated meaning of the combination and practice block. Error bars represent standard error of the mean.

Table 3

Outcomes of Omnibus Analyses of Variance for Experimental Tasks

Statistical test	df	F	MSE	p	η_p^2
Reading practice task					
Reading times in disambiguating region for subordinate and dominant items in Blocks 1–8					
Main effect of age	1,112	63.77	92,0195	<.001	.36
Main effect of meaning	1,112	132.68	36,688	<.001	.54
Main effect of practice block	7,784	479.73	110,252	<.001	.81
Age × Meaning	1,112	20.87	36,688	<.001	.16
Age × Practice block	7,784	3.09	110,252	.003	.03
Meaning × Practice block	7,784	38.61	21,532	<.001	.26
Three-way interaction	7,784	2.77	21,532	.008	.02
Reading times in spillover region for subordinate and dominant items in Blocks 1–8					
Main effect of age	1,112	70.25	433,901	<.001	.39
Main effect of meaning	1,112	41.21	13,361	<.001	.27
Main effect of practice block	7,784	330.80	43,954	<.001	.75
Age × Meaning	1,112	20.51	13,361	<.001	.16
Age × Practice block	7,784	6.60	43,954	<.001	.06
Meaning × Practice block	7,784	5.24	9,723	<.001	.05
Three-way interaction	7,784	1.49	9,723	.169	.01
Reading times in combination region in subordinate and dominant conditions in Blocks 1–8					
Main effect of age	1,112	77.74	465,953	<.001	.41
Main effect of meaning	1,112	2.99	16,222	.086	.03
Main effect of practice block	7,784	344.16	65,355	<.001	.75
Age × Meaning	1,112	1.09	16,222	>.10	
Age × Practice block	7,784	0.33	65,355	>.10	
Meaning × Practice block	7,784	1.33	11,001	>.10	
Three-way interaction	7,784	1.49	11,001	>.10	
Reading transfer task					
Reading times in disambiguating region					
Main effect of age	1,112	46.11	803,762	<.001	.29
Main effect of practice meaning	1,112	0.43	185,380	>.10	
Main effect of transfer meaning	1,112	13.95	154,964	<.001	.11
Age × Practice meaning	1,112	0.04	185,380	>.10	
Age × Transfer meaning	1,112	0.22	154,964	>.10	
Practice Meaning × Transfer meaning	1,112	93.32	164,027	<.001	.46
Three-way interaction	1,112	8.59	164,027	.004	.07
Alphabet arithmetic task					
Response times for correct responses					
Main effect of age	1,111	13.61	10,414,928	<.001	.11
Main effect of addend size	1,111	285.57	833,665	<.001	.72
Main effect of practice block	7,777	253.37	464,942	<.001	.70
Age × Addend size	1,111	19.59	833,665	<.001	.15
Age × Practice block	7,777	8.76	464,942	<.001	.07
Addend Size × Practice block	7,777	13.66	119,318	<.001	.11
Three-way interaction	7,777	1.90	119,318	.067	.02

Table 4

Outcomes of Paired Comparisons for Practice Blocks in Reading Practice Task and Alphabet Arithmetic Task

	Younger adults			Older adults		
Statistical test	$t(54)$	p	d	$t(58)^a$	p	d
Reading practice task						
Reading times in disambiguating region for subordinate versus dominant items						
Block 1	5.91	<.001	0.47	9.64	<.001	0.85
Block 2	3.17	.002	0.27	6.17	<.001	0.50
Block 3	2.48	.016	0.21	4.29	<.001	0.35
Block 4	0.27	.790	0.02	5.46	<.001	0.42
Block 5	1.16	.253	0.09	3.52	.001	0.31
Block 6	−0.84	.403	−0.07	1.49	.141	0.11
Block 7	1.75	.086	0.14	.30	.767	0.02
Block 8	1.22	.228	0.10	2.28	.026	0.15
Reading times in spillover region for subordinate versus dominant items						
Block 1	1.14	.262	0.09	4.26	<.001	0.39
Block 2	1.92	.060	0.16	3.63	.001	0.31
Block 3	−0.54	.590	−0.04	3.85	<.001	0.29
Block 4	2.24	.030	0.20	3.64	.001	0.34
Block 5	−0.46	.647	−0.04	1.03	.307	0.07
Block 6	−1.46	.150	−0.13	1.59	.118	0.11
Block 7	1.14	.260	0.11	1.02	.313	0.07
Block 8	−0.56	.577	−0.06	2.67	.010	0.14
Alphabet arithmetic						
Response times for correct responses for addend 4 versus addend 2 items						
Block 1	9.69	<.001	0.66	12.24	<.001	0.95
Block 2	8.75	<.001	0.73	12.73	<.001	1.06
Block 3	5.83	<.001	0.54	12.42	<.001	1.06
Block 4	5.51	<.001	0.44	11.71	<.001	1.02
Block 5	4.87	<.001	0.42	9.97	<.001	0.93
Block 6	4.81	<.001	0.37	10.91	<.001	0.93
Block 7	4.84	<.001	0.44	9.64	<.001	0.93
Block 8	2.99	.004	0.26	9.96	<.001	0.89

Note. Using the Bonferroni correction for multiple comparisons, the alpha level for significance in each set of comparisons is $p < .0063$.
^{a} $df = 57$ for alphabet arithmetic task.

In sum, the convergence of reading times in the subordinate and dominant conditions later in practice provides evidence that both age groups were shifting to retrieval. An alternative interpretation is that older adults were still engaging in computation in later blocks of practice but were completing reanalysis in the subsequent region (i.e., the *spillover region*) rather than in the disambiguating region. Examination of reading times in the spillover region largely rule out this possibility (see Figure 2; inferential statistics are reported in Tables 3-4). Although some spillover effects were evident early in practice, older adults' reading times for subordinate versus dominant items in the spillover region did not significantly differ in Blocks 5–7 (see Table 4).

Finally, reading times for the sentence region containing the combination (e.g., "the skunk mud") are reported in Table 5; outcomes of an omnibus 2 (age group) × 2 (subordinate vs. dominant condition) × 8 (practice block) mixed-factor ANOVA are reported in Table 3. As would be expected, the main effects of age group and practice block were significant; no other effects were significant.
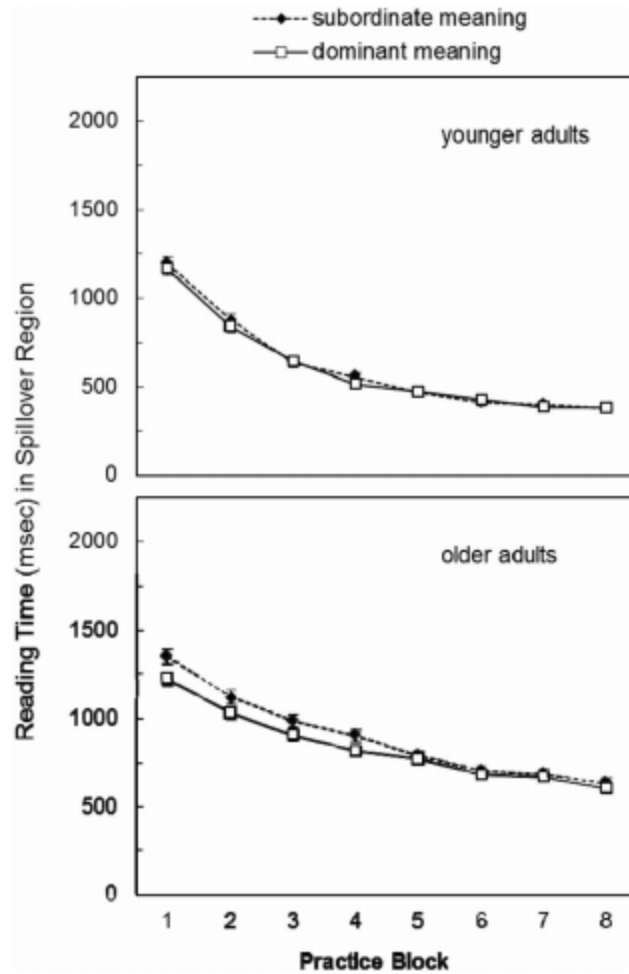
Figure 2. Mean reading time (in milliseconds) in the spillover region of target sentences in the reading practice task, for younger adults (top panel) and older adults (bottom panel), as a function of the disambiguated meaning of the combination and practice block. Error bars represent standard error of the mean.

Table 5
Reading Times (in Milliseconds) for Sentence Regions Containing Combination During Practice Blocks in Reading Practice Task

Block	Younger adults			Older adults		
	Subordinate	Dominant	d	Subordinate	Dominant	d
Block 1	1,281 (57)	1,273 (55)	0.02	1,549 (60)	1,533 (59)	0.03
Block 2	856 (31)	844 (32)	0.05	1,138 (44)	1,088 (40)	0.15
Block 3	626 (29)	643 (29)	−0.08	961 (32)	914 (38)	0.17
Block 4	529 (23)	530 (26)	0.00	851 (32)	801 (32)	0.20
Block 5	447 (19)	458 (21)	−0.08	742 (27)	777 (32)	−0.15
Block 6	411 (17)	394 (17)	0.14	676 (27)	675 (26)	0.00
Block 7	372 (15)	357 (13)	0.14	655 (30)	657 (28)	−0.01
Block 8	354 (14)	343 (13)	0.10	609 (32)	604 (28)	0.02

Reading Transfer Task

Outcomes of the reading transfer task provide even stronger evidence for older adults' preservation of memory-based automaticity during reading. Mean reading times in the disambiguating region of the transfer sentences in each condition are reported in Figure 3; outcomes of a 2 (age group) × 2 (meaning in practice) × 2 (meaning in transfer) ANOVA are reported in Table 3.

The key outcome concerns the significant two-way interaction between practice meaning and transfer meaning (see Table 3), which provides evidence that participants were retrieving interpretations stored during practice rather than computing interpretations during transfer. As noted in Table 3, the three-way interaction (between practice meaning, transfer meaning, and age) was significant. Follow-up tests revealed that the critical two-way interaction between practice meaning and transfer meaning (providing evidence for memory-based processing) was stronger for older adults—$\eta p^2 = .53$, $F(1, 58) = 64.22$, $MSE = 209833$, $p < .001$—than for younger adults— $\eta p^2 = .37$, $F(1, 54) = 31.25$, $MSE = 114828$, $p < .001$.

For both age groups, reading times on the subordinate meaning during transfer were significantly shorter when that item had been practiced with the subordinate meaning versus practiced with the dominant meaning—younger adult: $t(54) = 4.03$, $p < .001$, $d = .41$; older adult: $t(58) = 5.26$, $p < .001$, $d = .79$. Likewise, reading times on the dominant meaning during transfer were significantly shorter when that item had been practiced with the dominant meaning versus practiced with the subordinate meaning—younger adult: $t(54) = 3.79$, $p < .001$, $d = .42$; older adult: $t(58) = 5.74$, $p < .001$, $d = .82$. Taken together, these outcomes indicate that upon encounter of a combination in the transfer sentence, participants retrieved interpretations of that combination that were encoded during practice. When the retrieved meaning was consistent with the disambiguation later in the sentence (i.e., in the subordinate-subordinate and dominant-dominant conditions), no reanalysis was required. When the retrieved meaning was inconsistent with the disambiguation later in the sentence (i.e., in the dominant-subordinate and subordinate-dominant conditions), reanalysis was required to resolve the inconsistency and, thus, reading times were elevated in these two conditions relative to their counterpart conditions. In contrast, if older adults had reverted back to algorithmic processing during transfer (e.g., due to forgetting or to reduced access to the stored interpretations because of context shift), the dominant meaning of the combinations would have been generated regardless of the meaning during practice. If so, reanalysis would be required for subordinate transfer items and, thus, only a main effect of transfer meaning would have been expected, which clearly did not obtain.

One additional comparison is relevant for ruling out the possibility that older adults' interpretation of the combinations reverted back to algorithmic processing during transfer. For items that were practiced with the subordinate meaning, reading times were longer when the dominant meaning was presented in transfer than when the subordinate meaning was presented (i.e., the subordinate-dominant vs. subordinate-subordinate conditions)—younger adult: $t(54) = 2.30$, $p = .025$, $d = .24$; older adult: $t(58) = 3.75$, $p < .001$, $d = .57$. If interpretation reverted back to the algorithmic process (which generates the dominant meaning), reading times would have been *faster* when the disambiguating region contained the dominant meaning than when it contained the subordinate meaning.

Finally, we briefly report outcomes concerning the accuracy of participants' sensibility judgments. To revisit, the task was used to reduce the nominal similarity of the practice and transfer tasks and performance was examined primarily as an indicator of sufficient attention during the transfer task. Overall, judgment accuracy was relatively high for younger and older adults (M = 89% and 93%, respectively). However, accuracy was lower in the condition involving dominant meanings in practice with subordinate meanings in transfer (M = 74%) than in the other three conditions (all Ms = 92%), $F(1, 112) = 43.71$, $p < .001$, $\eta p^2 = .28$ (no interactions involving age group were significant). Although unexpected, this outcome is consistent with the possibility that individuals had particular difficulty resolving the inconsistency between practice and transfer meanings in the dominant-to-subordinate condition and/or rejected the subordinate meaning.

Alphabet Arithmetic Task

As a reminder, each AA item was presented four times in each practice block; values were averaged across trials in each block. Alphabet arithmetic data were excluded for one older adult with response times more than three standard deviations above group means in every block of practice (and >4 SDs in most blocks).

The percentage of correct responses on AA trials was consistently high in all eight blocks for both age groups (Ms = 95%–98%). For each participant, we computed mean response time for correct responses within each block for each addend condition.[4] Less than 1% of response times for individual trials were removed as outliers. Figure 4 reports mean response times for each block. Results of an omnibus 2 (age group) × 2 (addend size) × 8 (practice block) mixed-factor ANOVA are reported in Table 3. Below, we focus discussion on outcomes from paired comparisons that evaluate directional predictions motivated by previous research.

---

[4] All qualitative patterns reported below are the same when response times for all items are included.

□ subordinate meaning in practice
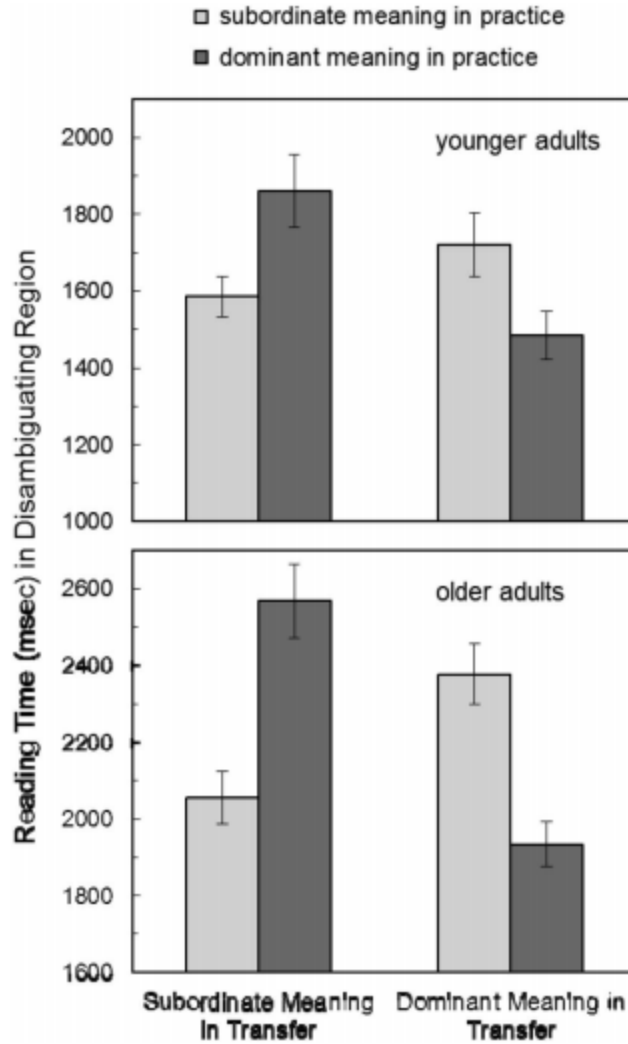■ dominant meaning in practice

Figure 3. Mean reading time (in milliseconds) in the disambiguating region of target sentences in the reading transfer task, for younger adults (top panel) and older adults (bottom panel), as a function of the disambiguated meaning of the combination during practice and during transfer. Error bars represent standard error of the mean.

As in prior research using similar tasks, both age groups showed practice effects, with significantly faster response times during the eighth versus first block in the addend 4 condition—younger adult: $t(54) = 12.68$, $p < .001$, $d = 2.18$; older adult: $t(57) = 13.89$, $p < .001$, $d = 1.66$—and in the addend 2 condition—younger adult: $t(54) = 13.27$, $p < .001$, $d = 1.97$; older adult: $t(57) = 15.93$, $p < .001$, $d = 1.51$. Of greatest interest, as is evident from the pattern shown in Figure 4, these practice effects apparently reflected almost no involvement of memory-based processing for older adults, with minimal convergence of response times for addend 2 and addend 4 items across practice (response time difference of 575 ms per addend in Block 1 vs. 404 per addend in Block 8). In contrast, although younger adults did not completely shift to memory-based processing by the end of practice (see Table 4), a stronger pattern of convergence is clearly evident (response time difference of 482 ms per addend in Block 1 vs. 133 per addend in Block 8).
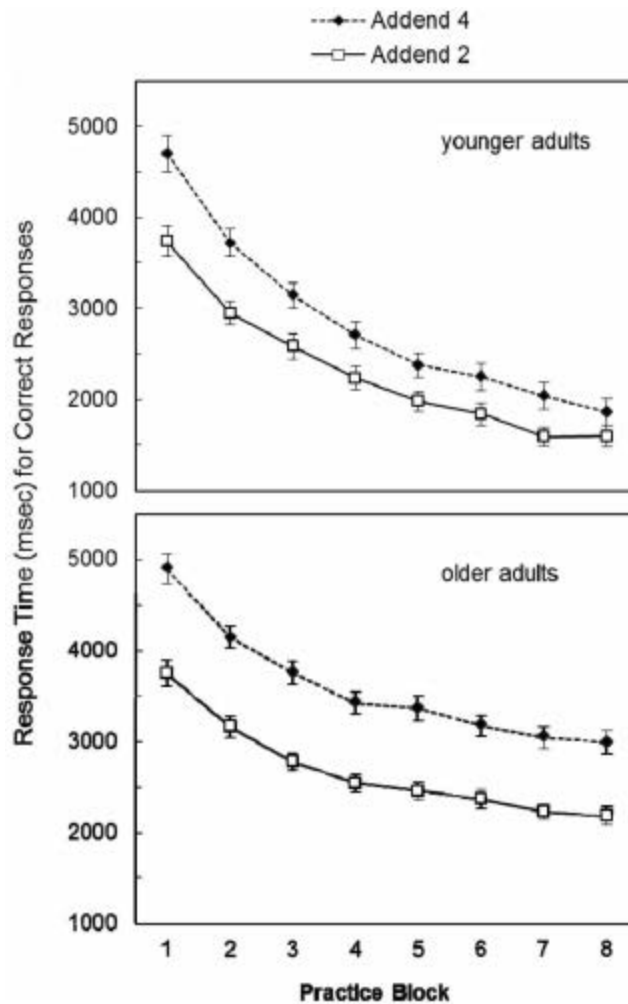
Figure 4. Mean response time (in milliseconds) for correct responses in the alphabet arithmetic task, for younger adults (top panel) and older adults (bottom panel), as a function of the size of the addend in each problem and practice block. As described in the text, each alphabet arithmetic item was presented 32 times, with four trials in each of eight blocks; each block value represents the mean response time across the four trials for items in that condition within that block. Error bars represent standard error of the mean.

Most important, the relatively minimal convergence in response times observed for older adults indicates that the rapid and robust shift to retrieval in the reading task was specific to reading and not task-general. To facilitate examination of the dissociation of age effects in the reading task versus the alphabet arithmetic task, Figure 5 illustrates the degree of convergence of response times across practice blocks in each task for each age group. Specifically, the top panel of Figure 5 reports the mean difference in reading times for subordinate versus dominant disambiguating regions in the reading task for each age group. The bottom panel of Figure 5 reports the mean difference in response times for addend 4 versus addend 2 items in the AA task for each age group. In all four cases, the data were best fit by an exponential function. To revisit, the magnitude of the difference in response times for conditions that involve more versus less complex computations indicates the degree to which memory-based processing is involved. For

the reading task, the difference in reading times approaches zero in both age groups, and the slope indicating rate of shift is actually somewhat steeper for older versus younger adults ($Z =$ 2.28, $p = .023$). In contrast, the rate of shift is much steeper for younger versus older adults in the AA task ($Z = 4.17, p < .001$).
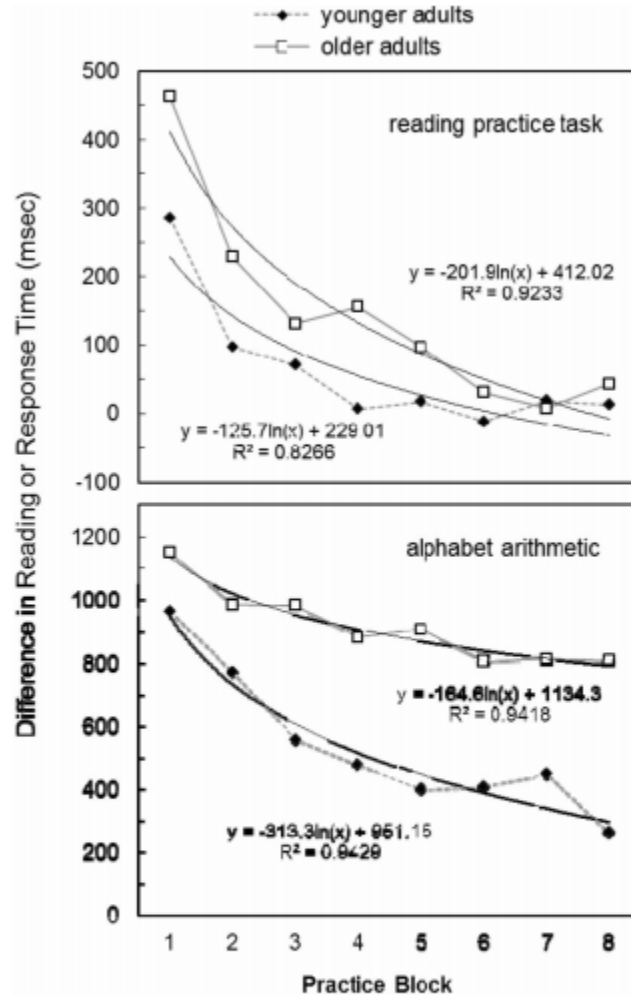

Figure 5. The top panel reports the mean difference between reading times for subordinate versus dominant items in each age group in each block of practice. The bottom panel reports the mean difference between response times for addend-4 versus addend-2 items in each age group in each block of practice. In all four cases, the mean values were better fit by a curvilinear (logarithmic) versus linear function; the curvilinear function and fit is reported for each condition.

DISCUSSION

In contrast to the sizable literature demonstrating deficits in memory-based automaticity for older adults, Rawson and Touron (2009) reported initial findings suggesting that older adults' memory-based automaticity is relatively preserved in a reading task. The current research provided more definitive evidence for robust memory-based automaticity for older adults during reading. During the reading practice task, older adults shifted to primary reliance on memory-based processing after relatively few repetitions of the unfamiliar noun phrases (as evidenced by

convergence of reading times in the subordinate and dominant conditions). This rapid shift to retrieval occurred despite somewhat more difficult practice conditions (larger set size and longer lags between repetitions) than in the original study by Rawson and Touron (2009).

Most important, older adults demonstrated robust memory-based processing when the noun phrases were presented in novel contexts in a nominally different transfer task—reading times differed substantially as a function of the match between meaning during practice and meaning during transfer. These outcomes support stronger claims about the robustness of memory-based automaticity in reading for older adults by establishing that memory-based interpretation was not contextually bound. Presumably, the slower reading times for mismatched conditions in transfer reflected the need for reanalysis when the information in the disambiguating region during transfer was inconsistent with the interpretation of the combination from practice that had been retrieved. Another possibility is that faster reading times in the matched conditions reflected repetition priming from repeated phrases that facilitated processing at the word level. This possibility is unlikely, given that repetition priming effects are typically small for word primes embedded in text (Levy & Kirsner, 1989; Oliphant, 1983; Rawson, 2007). The potential role of repetition priming raises the question of what exactly underlies priming—given that memory-based automaticity theorists assume that retrieval of previously stored interpretations also underlies repetition priming (Gupta & Cohen, 2002; Logan, 1990), this possibility does not pose interpretive difficulty for present purposes and is not mutually exclusive (i.e., the pattern of reading times in transfer may reflect both repetition priming and semantic reanalysis in the disambiguating region).

In contrast to older adults' relatively rapid shift to memory-based automaticity in reading (i.e., by Trial 6 during practice and then maintained into transfer), older adults showed minimal evidence for memory-based automaticity in alphabet arithmetic, aligning with prior research involving nonreading tasks (Hoyer et al., 2003; Jenkins & Hoyer, 2000; Rogers et al., 2000; Touron & Hertzog, 2004a, 2004b; Touron et al., 2004, 2007). Of course, the reading task used here differed from the alphabet arithmetic task (and other tasks used in prior automaticity research) in various aspects of the task and materials. However, at least two factors—set size and lag—can be ruled out as explanations for differential shifting (other factors are considered further below). Reading practice involved 24 repeated items, with trials for a given item separated by 23 other stories on average (as well as the secondary tasks inserted between blocks). Alphabet arithmetic only involved eight repeated items, and the four trials for a given item within each block were separated by only seven other items only requiring a few seconds each for response. Smaller set size and shorter lags would have favored the alphabet arithmetic task over the reading task, but the opposite pattern was observed.

Theoretical Implications

The overall pattern of outcomes has important implications for theories of memory-based automaticity in general and for specific theoretical accounts of age effects on memory-based automaticity. As described below, none of the existing theories as currently formulated are equipped to explain the current outcomes. Thus, these results will inform theory development that in turn will guide further research to evaluate plausible moderators of the involvement of memory-based processing during skill acquisition.

Memory-based theories of automaticity

The differential rate of shifting in reading tasks versus in nonreading tasks observed here and elsewhere is not easily accommodated by basic memory-based theories of automaticity (Logan, 1988; Palmeri, 1997; Rickard, 1997). To revisit, instance theory assumes that the shift from algorithm to retrieval is primarily based on (a) the relative speed of the algorithm versus retrieval, which in turn is largely determined by (b) the number of instances stored in memory. On average, older adults shifted to retrieval in the reading task by Trial 6 (i.e., with only five prior instances stored in memory), whereas they failed to shift by the end of AA practice even with 31 prior instances stored in memory. Thus, the differential shift in these two tasks cannot easily be attributed to differences in the number of stored instances. Differential shift also cannot easily be attributed to the relative speed of the algorithm versus retrieval in these two tasks, because algorithm speed was much slower in the AA task than in the reading task. For the AA task, response times in Block 1 (see Figure 4) provide an estimate of initial algorithm speed. For the reading task, reading times in Block 1 for the region of the target sentence containing the unfamiliar noun phrase (e.g., in the stories shown in Table 1, "The tiger chair" and "the skunk mud") provide an estimate of initial algorithm speed (as this is the region in which the meaning of the noun phrase would initially be processed via algorithm). For older adults, AA algorithm speed was 4,324 ms on average across the two addend conditions, whereas speed of the algorithm for interpreting noun phrases was 1,541 ms. Thus, retrieval should have been less competitive in the reading task than in the AA task.

Associative deficit hypothesis

As currently formulated, memory-based automaticity theories also do not afford a straightforward explanation for age differences, given that younger and older adults receive the same amount of practice with each item in each task and, thus, presumably had the same number of instances stored in memory. To account for age differences in memory-based automaticity, the associative deficit hypothesis (Naveh-Benjamin, 2000) states that older adults are less able to rely on retrieval due to poorer encoding of stimulus-response pairs during practice. This hypothesis is consistent with the voluminous literature documenting associative memory deficits for older adults (e.g., Cerella, Onyper, & Hoyer, 2006; Dunlosky, Hertzog, & Powell-Moman, 2005; Naveh-Benjamin, 2000; Naveh-Benjamin, Brav, & Levy, 2007; Old & Naveh-Benjamin, 2008), and older adults' retrieval use and response time improvements are related to associative memory ability (Rogers et al., 2000). Based on striking similarity between patterns of paired associate learning in an explicit memory task and shifts to retrieval in a skill training task, Cerella et al. (2006) concluded that "the age deficit in SK [skill training] can be accounted for by the age deficit in associative memory; no further explanation is needed" (p. 483).

However, the current results are not easily accommodated by the encoding deficiency account, because it does not provide a straightforward explanation for why older adults would be equipped to shift to memory-based processing in reading after very few trials but not in nonreading tasks after extensive amounts of practice. At a minimum, to be viable for explaining the overall pattern of outcomes, an encoding deficiency account would require further specification of the factors that would yield relatively intact associative memory in reading but not in other tasks.

One key factor may concern the richness of the representation afforded in reading versus in other tasks, which may affect the quality of encoding and, thus, the accessibility of instances to support memory-based processing. For example, higher-quality encoding during reading may be supported by the additional contextual information provided in the narrative that can be encoded along with the nominal stimulus (in the current task, the nominal stimulus refers to the novel noun-noun combinations) and the interpretation of that stimulus. In contrast, the nominal stimuli in other automaticity tasks are typically presented with no additional contextual information that would support encoding of stronger stimulus–response associations. One possible means for testing this account would involve presenting novel combinations and their intended meanings either embedded in a narrative or in isolation, followed by a memory-based criterion task. This account also predicts more generally that other tasks involving richer contextual information may also show relative preservation of memory-based automaticity for older adults (e.g., training through pictures or images).

Likewise, the nominal stimuli involved in the reading task versus other automaticity tasks likely differ in the nature of elaborative processing afforded by the nominal stimuli involved. In contrast to the stimulus-response pairs in alphabet arithmetic (and most other automaticity tasks used to date), the concreteness of the noun combinations and their meanings may allow older adults to create a rich semantic representation of the combination in memory, through visual imagery or some other form of elaboration. To explore this possibility, one direction for further research could involve presentation of the same nominal stimuli (e.g., *skunk mud*) as novel combinations in the reading task versus as noun pairs in the noun-pair look-up task (in brief, this task involves presenting a set of noun pairs in a table at the top of the screen, and on each trial either an intact or a mismatched noun pair is presented at the bottom of the screen for participants to decide whether that pair is included in the table). Equating the nominal stimuli in the two tasks would then permit examination of functional differences in the spontaneous use of elaborative processing strategies as a function of task, and/or differential sensitivity to instructional manipulations to affect the use of elaborative encoding in the two tasks (e.g., to the extent that older adults spontaneously use imagery to encode stimuli during reading but not during noun-pair look-up, instructions to use imagery would have a stronger effect on memory-based processing in the noun-pair look-up task than in the reading task, and a stronger effect on noun-pair look-up performance for older vs. younger adults).

Retrieval aversion hypothesis

Age differences in retrieval shifts have also been proposed to reflect strategic differences in older and younger adults' reliance on retrieval, even when associative memory is sufficient for use of memory-based processing (i.e., the *retrieval aversion hypothesis*, Touron & Hertzog, 2004a). Touron and Hertzog (2004a) used the noun-pair look-up task with a set of 20 noun pairs. Each item was presented for 100 trials, and participants reported which process they used on each trial (scan the table or retrieve the response). Despite the extensive amount of practice, reports of retrieval use were still substantially lower for older adults than for younger adults at the end of practice (around 55% vs. 95% of trials, respectively). Another group of participants periodically completed recognition trials in which they responded to noun pair probes without the look-up table present. Although memory performance was somewhat lower for older versus younger adults at the beginning of practice, memory performance did not differ and was

consistently near ceiling for the remainder of practice. Thus, despite having adequate item memory, older adults underutilized memory-based processing (see also Frank, Touron, & Hertzog, 2013; Hertzog & Touron, 2011; Touron & Hertzog 2004b, 2009; Touron et al., 2007).

As currently formulated, the retrieval aversion hypothesis is not equipped to account for the current outcomes. To accommodate the differential involvement of memory-based processing in reading versus nonreading tasks, a relatively straightforward modification of this account would assume that older adults are less averse to rely on retrieval in reading versus nonreading tasks. Some prior outcomes are consistent with this possibility. Rawson and Touron's (2009) Experiment 2 included a within-participant manipulation in which half of the noun phrases and definitions were pretrained, using a study-test procedure until each had been correctly recalled once. Older adults shifted rapidly for both trained and untrained noun phrases. This rapid and uniform shift to retrieval by older adults stands in direct contrast to the patterns obtained in comparable studies with other skill acquisition tasks. For example, Hines, Hertzog, and Touron (2012) used an extensive pretraining manipulation for the learning of noun pairs. When all noun pairs were pretrained to a high criterion, both young and older adults shifted to the memory retrieval strategy rapidly. However, when only half of the noun pairs were pretrained, older adults showed substantially less use of the memory strategy, even following considerable training with each noun pair. The retrieval aversion for the pretrained noun pairs in this condition implicates a general metacognitive influence on strategy choice, such as memory confidence. Taken together, these findings are consistent with the possibility that older adults are more likely to utilize memory retrieval strategies in reading, relative to other tasks.

Of course, this assumption begs the question of why older adults would be less averse to using retrieval during reading. One possibility is that retrieval aversion is a function of the demands for responding. Older adults may be more averse to relying on retrieval in tasks that require a speeded overt response that will be judged as objectively correct or incorrect (such as alphabet arithmetic and other novel laboratory tasks) versus in tasks that do not require an overt and/or speeded response (such as reading). Another possibility is that retrieval aversion is a function of general task familiarity, such that older adults may be more averse to relying on retrieval in unfamiliar tasks (such as alphabet arithmetic and other novel laboratory tasks) versus in familiar tasks like reading. It is also possible that older adults are more willing to utilize memory retrieval when it is supported by the rich context of a narrative passage. Although older adults are somewhat less precise when initially deriving the meaning of an unfamiliar word from its textual context (McGinnis & Zelinski, 2000, 2003), this effect was limited to the oldest subset of older adults (75 + years), and repeated exposure might provide sufficient support to adopt and rely on a meaning that has been explicitly provided, as in the present study. Investigating these and other factors that may systematically affect older adults' retrieval aversion is an important direction for future research.

In sum, the current research provides definitive evidence for relative preservation of memory-based automaticity in reading for older adults, which represents a significant departure from the modal finding in the vast majority of prior research involving nonreading tasks. These outcomes have important implications for theories of age effects in skill acquisition and for theories of memory-based automaticity more generally. This work also provides a useful launching point for

further investigation of age-related differences in the involvement of memory-based processing in cognitive skill acquisition.

REFERENCES

Brown, G. D. A. (1984). A frequency count of 190,000 words in the London-Lund Corpus of English Conversation. *Behavioural Research Methods, Instrumentation and Computers*, *16*, 502–532.

Brown, T. L., Gore, C. L., & Carr, T. H. (2002). Visual attention and word recognition in Stroop color naming: Is word recognition "automatic"?*Journal of Experimental Psychology: General*, *131*, 220–240.

Cerella, J., Onyper, S. V., & Hoyer, W. J. (2006). The associative-memory basis of cognitive skill learning: Adult age differences. *Psychology and Aging*, *21*, 483–498. 10.1037/0882-7974.21.3.483

Collins, W. M., & Levy, B. A. (2008). Developing fluent text processing with practice: Memorial influences on fluency and comprehension. *Canadian Psychology*, *49*, 133–139.

Cortina, J. M., & Nouri, H. (2000). *Effect size for ANOVA designs*. Thousand Oaks, CA: Sage.

Dunlosky, J., Hertzog, C., & Powell-Moman, A. (2005). The contribution of mediator-based deficiencies to age differences in associative learning. *Developmental Psychology*, *41*, 389–400. 10.1037/0012-1649.41.2.389

Estes, Z., & Glucksberg, S. (2000). Interactive property attribution in concept combination. *Memory & Cognition*, *28*, 28–34. 10.3758/BF03211572

Fisk, A. D., Hertzog, C., Lee, M. D., Rogers, W. A., & Anderson-Garlach, M. (1994). Long-term retention of skilled visual search: Do young adults retain more than old adults?*Psychology and Aging*, *9*, 206–215. 10.1037/0882-7974.9.2.206

Frank, D. J., Touron, D. R., & Hertzog, C. (2013). Age differences in strategy shift: Retrieval avoidance or general shift reluctance?*Psychology and Aging*, *28*, 778–788. 10.1037/a0030473

Gagné, C. L., & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 71–87. 10.1037/0278-7393.23.1.71

Gerrig, R. J., & Bortfeld, H. (1999). Sense creation in and out of discourse contexts. *Journal of Memory and Language*, *41*, 457–468. 10.1006/jmla.1999.2656

Grant, S. C., & Logan, G. D. (1993). The loss of repetition priming and automaticity over time as a function of degree of initial learning. *Memory & Cognition*, *21*, 611–618. 10.3758/BF03197193

Gupta, P., & Cohen, N. J. (2002). Theoretical and computational analysis of skill learning, repetition priming, and procedural memory. *Psychological Review*, *109*, 401–448.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, *11*, 194–205. 10.1162/089892999563328

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, *108*, 356–388. 10.1037/0096-3445.108.3.356

Hertzog, C., & Touron, D. R. (2011). Age differences in memory retrieval shift: Governed by feeling-of-knowing?*Psychology and Aging*, *26*, 647–660. 10.1037/a0021875

Hertzog, C., Touron, D. R., & Hines, J. C. (2007). Does a time-monitoring deficit influence older adults' delayed retrieval shift during skill acquisition?*Psychology and Aging*, *22*, 607–624. 10.1037/0882-7974.22.3.607

Hines, J., Hertzog, C., & Touron, D. (2012). A prelearning manipulation falsifies a pure associational deficit account of retrieval shift during skill acquisition. *Aging, Neuropsychology, and Cognition*, *19*, 449–478.

Hoyer, W. J., Cerella, J., & Onyper, S. V. (2003). Item learning in cognitive skill training: Effects of item difficulty. *Memory & Cognition*, *31*, 1260–1270. 10.3758/BF03195809

Jenkins, L., & Hoyer, W. J. (2000). Instance-based automaticity and aging: Acquisition, reacquisition, and long-term retention. *Psychology and Aging*, *15*, 551–565. 10.1037/0882-7974.15.3.551

Klapp, S. T., Boches, C. A., Trabert, M. L., & Logan, G. D. (1991). Automatizing alphabet arithmetic: II. Are there practice effects after automaticity is achieved?*Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 196–209. 10.1037/0278-7393.17.2.196

Lamson, N., & Rogers, W. A. (2008). Assessing age-related patterns in strategy selection on a mathematical problem-solving task. *The Journals of Gerontology: Series B. Psychological Sciences and Social Sciences*, *63*, 146–155. 10.1093/geronb/63.3.P146

Lassaline, M. E., & Logan, G. D. (1993). Memory-based automaticity in the discrimination of visual numerosity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 561–581. 10.1037/0278-7393.19.3.561

Levy, B. A., & Burns, K. I. (1990). Reprocessing text: Contributions from conceptually driven processes. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *44*, 465–482. 10.1037/h0084265

Levy, B. A., & Kirsner, K. (1989). Reprocessing text: Indirect measures of word and message level processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 407–417. 10.1037/0278-7393.15.3.407

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527. 10.1037/0033-295X.95.4.492

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms?*Cognitive Psychology*, *22*, 1–35.

Logan, G. D., & Klapp, S. T. (1991). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity?*Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 179–195. 10.1037/0278-7393.17.2.179

McGinnis, D., & Zelinski, E. M. (2000). Understanding unfamiliar words: The influence of processing resources, vocabulary knowledge, and age. *Psychology and Aging*, *15*, 335–350. 10.1037/0882-7974.15.2.335

McGinnis, D., & Zelinski, E. M. (2003). Understanding unfamiliar words in young, young-old, and old-old adults: Inferential processing and the abstraction-deficit hypothesis. *Psychology and Aging*, *18*, 497–509. 10.1037/0882-7974.18.3.497

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, *99*, 440–466. 10.1037/0033-295X.99.3.440

Middleton, E. L., Rawson, K. A., & Wisniewski, E. J. (2011). How do we process novel conceptual combinations in context?*The Quarterly Journal of Experimental Psychology*, *64*, 807–822. 10.1080/17470218.2010.520414

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1170–1187. 10.1037/0278-7393.26.5.1170

Naveh-Benjamin, M., Brav, T. K., & Levy, O. (2007). The associative memory deficit of older adults: The role of strategy utilization. *Psychology and Aging*, *22*, 202–208. 10.1037/0882-7974.22.1.202

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, *23*, 104–118. 10.1037/0882-7974.23.1.104

Oliphant, G. W. (1983). Repetition and recency effects in word recognition. *Australian Journal of Psychology*, *35*, 393–403. 10.1080/00049538308258751

Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 324–354. 10.1037/0278-7393.23.2.324

Pashler, H., & Harris, C. R. (2012). Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science*, *7*, 531–536. 10.1177/1745691612463401

Rawson, K. A. (2004). Exploring automaticity in text processing: Syntactic ambiguity as a test case. *Cognitive Psychology*, *49*, 333–369. 10.1016/j.cogpsych.2004.04.001

Rawson, K. A. (2007). Testing the shared resource assumption in theories of text processing. *Cognitive Psychology*, *54*, 155–183. 10.1016/j.cogpsych.2006.06.002

Rawson, K. A. (2010). Defining and investigating automaticity in reading comprehension. In B. H.Ross (Ed.), *Psychology of learning and motivation* (Vol. *52*, pp. 185–230). Burlington, VT: Academic Press. 10.1016/S0079-7421(10)52005-X

Rawson, K. A., & Middleton, E. L. (2009). Memory-based processing as a mechanism of automaticity in text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 353–370. 10.1037/a0014733

Rawson, K. A., & Touron, D. R. (2009). Age differences and similarities in the shift from computation to retrieval during reading comprehension. *Psychology and Aging*, *24*, 423–437. 10.1037/a0016044

Rickard, T. C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General*, *126*, 288–311. 10.1037/0096-3445.126.3.288

Rickard, T. C., Lau, J. S-H., & Pashler, H. (2008). Spacing and the transition from calculation to retrieval. *Psychonomic Bulletin & Review*, *15*, 656–661. 10.3758/PBR.15.3.656

Roediger, H. L., III. (2012). Psychology's woes and a partial cure: The value of replication. *APS Observer*, *25*, 9, 27–29.

Rogers, W. A., Hertzog, C., & Fisk, A. D. (2000). An individual differences analysis of ability and strategy influences: Age-related differences in associative learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 359–394. 10.1037/0278-7393.26.2.359

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social science. *Review of General Psychology*, *13*, 90–100. 10.1037/a0015108

Schneider, W., Dumais, S. T., & Shiffrin, R. M. (1984). Automatic and control processing and attention. In R.Parasuraman & R.Davies (Eds.), *Varieties of attention* (pp. 1–27). New York, NY: Academic Press.

Stine-Morrow, E. A. L., Gagne, D. D., Morrow, D. G., & DeWall, B. H. (2004). Age differences in rereading. *Memory & Cognition*, *32*, 696–710. 10.3758/BF03195860

Touron, D. R. (2006). Are item-level strategy shifts abrupt and collective? Age differences in cognitive skill acquisition. *Psychonomic Bulletin & Review*, *13*, 781–786. 10.3758/BF03193997

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. 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. 10.3758/BF03196860

Touron, D. R., & Hertzog, C. (2009). Age differences in strategic behavior during a computation-based skill acquisition task. *Psychology and Aging*, *24*, 574–585.

Touron, D. R., Hoyer, W. J., & Cerella, J. (2004). Cognitive skill learning: Age-related differences in strategy shifts and speed of component operations. *Psychology and Aging*, *19*, 565–580. 10.1037/0882-7974.19.4.565

Touron, D. R., Swaim, E. T., & Hertzog, C. (2007). Moderation of older adults' retrieval reluctance through task instructions and monetary incentives. *The Journals of Gerontology: Series B. Psychological Sciences and Social Sciences*, *62*, 149–155. 10.1093/geronb/62.3.P149

Walczyk, J. J. (2000). The interplay between automatic and control processes in reading. *Reading Research Quarterly*, *35*, 554–566. 10.1598/RRQ.35.4.7

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale—Revised*. New York, NY: Psychological Corporation.

Wilkins, N. J., & Rawson, K. A. (2010). Loss of cognitive skill across delays: Constraints for theories of cognitive skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1134–1149. 10.1037/a0019998

Wilkins, N. J., & Rawson, K. A. (2013). Why does lag affect the durability of memory-based automaticity: Loss of memory strength or interference? *Acta Psychologica*, *144*, 390–396. 10.1016/j.actpsy.2013.07.021

Wilkins, N. J., & Rawson, K. A. (2014, May). *Does set size affect memory-based automaticity?* Paper presented at the 86th annual meeting of the Midwestern Psychological Association, Chicago, IL.

Wisniewski, E. J. (1997). When concepts combine. *Psychonomic Bulletin & Review*, *4*, 167–183. 10.3758/BF03209392

Wisniewski, E. J., & Middleton, E. L. (2002). Of bucket bowls and coffee cup bowls: Spatial alignment in conceptual combination. *Journal of Memory and Language*, *46*, 1–23. 10.1006/jmla.2001.2801

Zachary, R. (1986). *Shipley Institute of Living Scale Revised manual*. Los Angeles: Western Psychological Services.