Metacognition for Strategy Selection during Arithmetic

Problem Solving in Young and Older Adults

Marie Geurten[1] & Patrick Lemaire[2,3]

[1] Psychology and Neuroscience of Cognition Unit, University of Liège, Belgium

[2] Aix-Marseille University, CNRS, LPC Marseille, France

[3] Institut Universitaire de France

Correspondence concerning this article should be addressed to Marie Geurten, University of Liège, B33 Trifacultaire – Quartier Agora, Place des Orateurs 1, 4000 Liège – Belgium; E-mail: mgeurten@ulg.ac.be; Phone number: +32 4 366 59 43

Patrick Lemaire, Aix-Marseille University & CNRS, Institut Universitaire de France, Case D - Place Victor Hugo 3, 13331 Marseille – France; E-mail: patrick.lemaire@univ-amu.fr

Acknowledgments: This work was supported by the Agence Nationale de la Recherche under under Grant # ANR-17-CE28-17-0003-01 to PL and by a Grant from the Marie-Curie CoFund Program to MG. We would like to thank Fleur Brun and Laura Bral for their help in data collection. We have no conflict of interest to declare.

Abstract

We examined participants' strategy choices and metacognitive judgments during arithmetic problem solving. Metacognitive judgments were collected either prospectively or retrospectively. We tested whether metacognitive judgments are related to strategy choices on the current problems and on the immediately following problems, and age-related differences in relations between metacognition and strategy choices. Data showed that both young and older adults were able to make accurate retrospective, but not prospective, judgments. Moreover, the accuracy of retrospective judgments was comparable in young and older adults when participants had to select and execute the better strategy. Metacognitive accuracy was even higher in older adults when participants had to only select the better strategy. Finally, low-confidence judgments on current items were more frequently followed by better strategy selection on immediately succeeding items than high-confidence judgments in both young and older adults. Implications of these findings to further our understanding of age-related differences and similarities in adults' metacognitive monitoring and metacognitive regulation for strategy selection in the context of arithmetic problem solving are discussed.

Keywords: Metacognition; Monitoring; Regulation; Aging; Strategy Selection; Arithmetic

Introduction

Metacognition refers to processes used to monitor and regulate our mental activities (Nelson, 1996). Four decades of research – mainly, but not only, in the memory domain– have established that the influence of metacognitive processes on cognitive performance occurs via implementation of effective strategies. Specifically, various studies have shown that participants who accurately monitor their internal states are more likely to implement appropriate strategies to regulate or control their performance, for instance, by allocating more time to more difficult items or by selecting the items on which they want to be tested (see Dunlosky & Metcalfe, 2009, for an overview). Despite a number of findings consistent with influence of metacognition on cognitive performance, surprisingly little research has been carried out to investigate how crucial metacognitive monitoring (i.e., the ability to evaluate one's own performance) and control (i.e., the ability to regulate one's own performance) are involved during strategy selection (for exceptions in the memory domains, see Kimball, Smith, & Muntean, 2012; Hines, Touron, & Hertzog, 2009; Touron, Hertzog, & Hines, 2007). Thus, we know little about how the accuracy of participants' estimates impacts their strategy choices, and how this changes with age. The present study contributes to these issues by examining age-related differences in metacognitive monitoring and metacognitive regulation for strategy selection in arithmetic problem solving for several reasons. This domain is relevant for everyday life (e.g., managing budgets, etc.). It is also a domain where selecting the better strategy leads to significant improvement on performance. Interestingly, many previous studies found that older adults are at a disadvantage relative to young adults in arithmetic problem solving (see Uittenhove & Lemaire, 2015, for a review). Consequently, studying age-related differences in metacognitive skills for strategy selection in arithmetic allows to ensure that potential differences observed between age groups at the metacognitive level, is not due to age-related differences at the level of cognitive operations. Before outlining the logic of the present work, we first review previous findings on metacognition and strategy selection and, then, on metacognition, strategy selection, and aging.

Metacognition and Strategy Selection

A plethora of studies have shown that people use a variety of strategies to accomplish cognitive tasks and that participants' performance as well as age-related changes in cognitive performance depend on which strategies are used on each item (Siegler, 2007). In arithmetic, determiners of participants' performance like which strategy they use, which type of problems they solve, and under what conditions have been studied in great detail (see Kadosh & Dowker, 2015 for an overview). However, the role of metacognitive processes on participants' performance, and age-related differences therein, has been much less studied. Thus, previous research has not determined if young and older participants are able to monitor whether they will select the better strategy on a forthcoming item (i.e., prospective judgment) or to estimate their level of confidence associated with a just-selected strategy (i.e., retrospective judgment). Yet, being able to estimate past and future strategy selection on each problem could be an essential prerequisite for people to efficiently regulate their strategic behaviors (e.g., Lovett & Schunn, 1999; Siegler & Araya, 2005). Interestingly, some data suggest that both young and older adults can use better strategy judgments to change strategy while executing an already-selected strategy (e.g., Ardiale & Lemaire, 2012; Ardiale & Lemaire, 2013).

Ardiale and Lemaire (2012, 2013) asked young and older adults to estimate products of two-digit multiplication problems like 37 x 64. Problems were displayed with a cue indicating which of two rounding strategies to use. After executing this cued strategy for one second (too short to fully complete strategy execution), participants could choose whether to change strategy (or not) if they judged that the cued strategy was not the best strategy for this item (i.e., the strategy that yields the closest estimate to correct product). The authors found that both young and older adults were able to interrupt execution of strategy and switch strategy when the cued strategy was not the best one. One important limitation of Ardiale and Lemaire's work is that participants' better strategy judgments were not based on their own strategy selection but on strategies selected by the experimenter. Furthermore, better strategy judgements were not directly assessed, as participants were not asked explicitly whether the cued strategy was the better or poorer strategy. Although rates of strategy switching may be a direct consequence of strategy judgments, participants may have switched strategies on some or many items without systematically evaluating whether the cued strategy was the better or poorer strategy. To date, it is unknown how young and older adults monitor their chances of selecting the better strategy on each item, how such strategy monitoring changes with age during adulthood, and whether participants are able to use the result of strategy monitoring to regulate subsequent strategy selection (i.e., metacognitive regulation or control). Indeed, most previous studies exploring interactions between metacognitive monitoring and regulation processes on an item-by-item basis examined whether low levels of confidence in one's own performance on an item lead to the implementation of appropriate strategies on the same item (e.g., an increase in study time; Metcalfe & Kornell, 2003; Metcalfe & Kornell, 2005; Nelson & Leonesio, 1988; Son & Metcalfe, 2000; Thiede & Dunlosky, 1999). No studies have tested whether low confidence in a selected strategy leads participants to adjust their strategy selection on an upcoming item.

The issues of strategy monitoring and strategy regulation are theoretically important for models of strategy selection. Some assumptions made by computational models of strategy selection are consistent with the hypothesis that being able to introspect on how easy it would be to select the better strategy or on the level of uncertainty associated with the selected strategy increases the likelihood of choosing the best strategy on each item. Generally, computational models propose that choosing among multiple strategies involves associative mechanisms such as activating the relative costs/benefits of each strategy and selecting the strategy that works best for a given problem on the basis of problem and strategy characteristics (Lovett & Anderson, 1996; Lovett & Schunn, 1999; Neches, 1987; Rieskamp & Otto, 2006; Siegler & Araya, 2005; Siegler and Shipley, 1995). In addition to associative mechanisms, two of the existing computational models assume that strategy choices involve metacognitive mechanisms. Specifically, in the Lovett and Schunn's (1999) Represent, Construct, Choose, Learn (RCCL) model, the metacognitive system enables participants to interrupt a strategy mid-execution if they estimate that the current strategy is not the best one or if it is inappropriate. In Siegler and Araya's (2005) Strategy, Choice, and Discovery Simulation* (SCADS*), the metacognitive system is crucial to create or discover new strategies. In sum, models of strategies include assumptions relative to metacognitive processes to evaluate the strategies once selected and, possibly to interrupt strategies mid-execution to switch for a better strategy (RCCL) or to create and discover new legitimate strategies (SCADS*). Consistent with this view, Geurten and Lemaire (2017) have recently reported some experimental data suggesting that metacognitive processes could be involved very early in the strategy selection processes, when participants try to select the better strategy on each problem. Specifically, they have shown that participants are able to decide whether a cued strategy is the better to solve each problem and to regulate their response decisions accordingly. However, in this experiment, participants judged the relevance of the cued strategy; they did not estimate their own strategy selection performance.

In this context, the first goal of the present study was to determine whether young and older adults are able to make accurate metacognitive judgments on their own strategy selection and whether they rely on these judgments to guide subsequent strategy choices. These data could help us determine whether and how metacognitive processes influence people's strategy selection.

Metacognition and Strategy Selection in Aging.

Another important reason to focus on the accuracy of people's metacognitive judgments for strategy selection is that recent research has suggested that some aspects of metacognition remain intact with advancing age (see Castel, Middlebrooks, & McGillivray, 2016, Hertzog, 2016, for overviews). For instance, Hertzog, Kidder, Powell-Moman, and Dunlosky (2002) presented participants with either related or unrelated pairs of words and asked young and older adults to estimate whether they thought they would be able to recall these pairs of words on an upcoming recall test. Results indicated that older adults' metacognitive judgments were closer to their actual memory performance whereas young adults were less accurate. Using a similar paradigm except that participants had the opportunity to restudy all items after making their metacognitive judgments, Hines, Touron, and Hertzog (2009) found that both young and older adults studied items that had been assigned lower confidence judgments longer than items that had been assigned higher confidence judgments, suggesting no age deficits in using memory monitoring to regulate learning. (see also, Eakin, Hertzog, & Harris, 2014; Kuhlmann & Touron, 2011; Price, Hertzog, & Dunlosky, 2010; Price & Murray, 2012; Robinson, Hertzog, & Dunlosky, 2006).These findings argue for a relatively spared ability to monitor and regulate mental operations across the adult life span. This sparing suggests that older adults may potentially be able to use their good metacognitive skills to partially overcome or compensate for negative effects of aging on cognitive performance by investing limited resources into areas that yield optimal return (Baltes & Baltes, 1990; Hertzog & Dunlosky, 2011).

However, results of research that has examined metacognitive monitoring and metacognitive regulation in aging are not as consistent as they appear. Several studies found age-related changes in metacognitive monitoring and control (e.g., Dunlosky & Connor, 1997; Pansky, Goldsmith, Koriat, & Pearlman-Avnion, 2009; Souchay & Isingrini, 2004; see

Dodson, 2017 for a review). For instance, Bunnell, Baken, and Richards-Ward (1999) found that older adults' metacognitive judgments prior to list presentation or after recall reflected overconfidence compared to actual working memory performance. In contrast, young adults showed greater metacognitive accuracy. Regarding metacognitive regulation, Souchay and Isingrini (2004) found that older adults did not allocate their study time as efficiently as young adults during a self-paced learning task. Similar age-related differences were found in a study examining whether participants accurately selected items for restudy on the basis of their previous metacognitive judgments (Tullis & Benjamin, 2012).

Contradictory findings regarding age-related differences and similarities in metacognition further highlight the importance of studying metacognitive monitoring and regulation processes to better understand which aspects of metacognition are spared and which aspects are impaired during aging. One potential explanation for inconsistent findings regarding age-related differences in metacognition is that older and young adults may differentially allocate their cognitive resources to monitoring and performance, depending on the task. Indeed, it is now well established that normal aging is frequently characterized by decreased attentional and executive resources (see Glisky, 2007, for an overview). In this context, it is possible that the accuracy of older adults' metacognitive judgments and regulation depends on the amount of resources required to accomplish cognitive tasks, with some tasks requiring more resources than others. Spared metacognitive skills with age would be found for less resource-consuming tasks, and age-related changes for more resource-consuming tasks. Consistent with this possibility, Sacher, Isingrini, and Taconnat (2013) found that when young adults made metacognitive judgments in a divided attention situation, their pattern of responses mimicked the decrease in metacognitive accuracy observed in older adults. Such findings suggests that the availability of cognitive resources is an important determinant of accurate metacognitive judgments.

As these studies were conducted neither in the arithmetic domain nor on strategy selection, the second goal of our study was to examine age-related differences in the accuracy of metacognitive monitoring and metacognitive regulation for strategy selection in arithmetic when the better strategy selection task involves different levels of difficulty. This was expected to enable us to determine not only whether there are age-related changes in metacognitive monitoring and regulation for strategy selection, but also to inform us on factors that possibly influence variations in the accuracy of metacognitive processes with age.

Overview of the Present Study

The primary goals of this study were to (a) document age-related differences in metacognition for strategy selection while young and older adults are asked to solve arithmetic problems, and (b) determine whether these age-related differences are influenced by the amount of cognitive resources that are needed to complete the task. Specifically, all individuals were given three different tasks, a selection-execution task (i.e., participants were given arithmetic problems and had to select the better strategy among two available strategies, and keep this selected strategy in mind while executing it), a selection-only task (i.e., participants were given arithmetic problems, and were asked to only select the better strategy without executing it), and an execution-only task (i.e., participants were given arithmetic problems together with the better strategy and had to execute the cued strategy). Previous research in this arithmetic task has revealed that selecting and executing the better strategy is usually harder and requires more cognitive resources than just selecting the better strategy (e.g., Lemaire, Arnaud, & Lecacheur, 2004).

Empirical works on strategies in arithmetic have shown that strategy selection and strategy execution are influenced by participants, stimulus, and situation characteristics (Siegler, 2007). These factors act individually and in interaction with each other. For example, Lemaire et al. (2004) asked young and older adults to provide estimates of two-digit multiplication problems (e.g., 43x38) with a rounding-down strategy (doing 40x30=1200) or a rounding-up strategy (doing 50x40=2000), under different levels of speed/accuracy pressures. The authors found that older adults executed the rounding-down strategy more slowly under accuracy-pressure conditions than under no-pressure conditions, especially when they solved easy problems. Young adults were less affected than older adults by the time-pressure manipulation. Such findings show that young and older adults are differently influenced by problem and task characteristics. For this reason, the set of problems used in the present experiment was carefully selected so as to explore whether problem features that are known to influence strategy selection and strategy execution influence metacognitive judgments and regulation. The two main problem characteristics tested here were first which strategy is better for a given problem (half the problems were best estimated with one strategy, the other problems with another strategy) and the size of unit digits (i.e., problems with both unit digits vs. one unit digit smaller or larger than 5).

In the selection-execution and in the selection-only tasks, metacognition was assessed using both Ease of Selection judgments (EoS) and Retrospective Confidence Judgments (RCJ). Specifically, young and older adults were asked to judge on a trial-by-trial basis the ease with which they would select the better strategy (i.e., prospective judgments) or the level of confidence associated with their having selected the better strategy (i.e., retrospective judgments). The EoS judgments were similar to other classical prospective judgments, such as Ease-of-Learning, except that participants were asked to judge their future strategy selection instead of judging their future performance. This procedure was used because several studies have shown that some aspects of prospective metacognitive judgments (i.e., prediction of future performance) differ from retrospective metacognitive judgments (i.e., estimation of the accuracy of past responses). For instance, Kelemen, Frost, and Weaver (2000) showed no significant correlations between these two types of judgments for the same task. Moreover, prospective judgments are generally less accurate than retrospective judgments in both young and older adults (von der Linden, Löffler, & Schneider, 2016). For instance, Siedlecka, Paulewicz, and Wierzchoń (2016) compared prospective and retrospective judgments of performance while participants solved anagrams. Participants rated their confidence that a particular word was the solution before or after giving their responses. These authors found more accurate judgments when participants gave these judgments after relative to before the responses. This difference is likely due to retrospective judgments being based on information supporting a particular choice due to continuous tracking of evidence such as response fluency (Alter & Oppenheimer, 2009; Fleming, Massoni, Gajdos, & Vergnaud, 2016; Siedlecka et al., 2016). Indeed, Siedlecka et al.'s results suggest that the simple act of making a response provides an additional cue to improve accuracy of metacognitive judgments. Moreover, some data show that prospective judgments are more likely based on recent history of performance and, thus, can be biased when the level of difficulty of a task varies from one item to the next (e.g., Fleming et al., 2016; Geurten & Meulemans, 2017).

To assess metacognitive monitoring, different analytical approaches are available. One of the most common approaches consists in computing a score of metacognitive accuracy (or resolution) reflecting how well one discriminates between correct and incorrect responses when making different types of metacognitive judgments. From a statistical perspective, several types of index can be computed (e.g., Murayama, Sakaki, Yan, & Smith, 2014). Each of them has unique strengths and weaknesses (for a detailed description of the characteristics of these metacognitive metrics, see Fleming & Lau, 2014). Here, we used the A'_{ROC} index which provides a bias-free measure of metacognition (i.e., this measure is not influenced by the tendency to use higher or lower confidence ratings; Galvin, Podd, Drga, & Whitmore, 2003).

A similar approach was adopted to assess metacognitive regulation. To determine whether participants use the result of their metacognitive monitoring to improve their subsequent strategy selection, the non-parametric *phi* (*φ)* coefficients was computed to estimate the accuracy between metacognitive judgments and rates of subsequent better strategy selection.

Regarding metacognitive monitoring, the hypothesis that strategy selection involves metacognitive processes predicts that metacognitive accuracy should be higher than chance for both the prospective and the retrospective judgments. Such findings would indicate that young and older participants can introspect on their strategy choices. Based on previous findings in the metacognition literature, two alternative developmental scenarios were possible. First, metacognitive monitoring, as involved in arithmetic strategy selection, changes with adults' age and varies with different task parameters (e.g., Sacher et al., 2013). Such a possibility might arise if metacognitive processes depend on available resources, and availability of resources varies with participants' age and experimental tasks/conditions. This hypothesis predicts that the accuracy of metacognitive judgments is influenced by participants' age (i.e., it should be lower in older adults), by problem features (i.e., it should be lower on most difficult problems), and by selection versus selection-execution tasks (i.e., it should be lower in the selection-execution task). Alternatively, metacognitive monitoring is an age-invariant, general skill (e.g., Hertzog et al., 2002). That is, it should be stable with age during adulthood.

Regarding metacognitive regulation, no studies have previously examined whether people can accurately use metacognitive judgment on one item to regulate their performance on the following item. However, previous findings examining how participants rely on the result of their metacognitive evaluation to control their performance for the same item suggest that the same developmental scenarios as for metacognitive monitoring are possible. First, it is possible that metacognitive regulation changes with adults' age and varies with different task parameters (e.g., Souchay & Isingrini, 2004). Second, metacognitive monitoring is an age-invariant skill that is stable with age during the course of adulthood (e.g., Hines et al., 2009).

In sum, this study had two main goals. First, we examined whether young and older adults are able to make accurate metacognitive judgments on strategy selection. We also investigated whether participant (young or older), problem (homogeneous or heterogeneous unit digits; rounding-down or rounding-up), and task (selection-only, selection-execution, or execution-only) characteristics (i.e., factors that are known to affect strategy choices; see Lemaire, 2016, for an overview) influence the accuracy of participants' metacognitive judgments for strategy selection. Second, we determined whether participants are able to use the result of their metacognitive monitoring to regulate their subsequent strategy choices and whether the accuracy of these metacognitive regulatory processes can be influenced by participant and task characteristics.

Method

Participants

Sixty six French-speaking participants were tested. There were two age groups: 37 healthy young adults (29 females; mean age = 21.14 years; SD = 2.01; age range = 18–29) and 29 healthy older adults (18 females; mean age = 72.99 years; SD = 6.77; age range = 64–88). Young adults were undergraduates from Aix-Marseille University who voluntarily participated. Older adults were volunteers recruited from the community of Liege in Belgium. No older adults were excluded as they all had scores larger than 27 (M = 29.1) in the Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975). Data collection stopped when the number of participants was sufficient to reach a predicted power of .80 (alpha = .05, beta = .20) for a triple Age x Condition x Problem Type interaction (medium effect size, *f* = .25; Cohen, 1988) with Age as the only between-participant factor. The required effect size was determined on the basis of similar research in laboratory settings examining strategy selection in young and older adults (Lemaire et al., 2004). Information about participants' sex, age, verbal fluency, and arithmetic fluency was collected at the end of the experiment. The latter two variables were assessed using the French version of the Mill-Hill Vocabulary Scale (Deltour, 1993) and the addition and subtraction-multiplication subtests of the French Kit (French, Ekstrom, & Price, 1963). As often found, older participants' arithmetic fluency (139 vs. 108; $F(1,64) = 13.45$, $p < .001$, $\eta^2_p = .17$) was larger than that of young adults. No differences were found for verbal fluency (22.2 vs. 20.5; $F < 2$).

Materials

The stimuli were 32 multiplication problems presented in a standard form (i.e., *a x b*) with the operands *a* and *b* being two-digit numbers. The same problems were used in each phase of the experiment in order to allow direct comparisons of participants' metacognitive accuracy between conditions. Based on the size of unit digits, half the problems ($N = 16$) were categorized as homogeneous and half as heterogeneous. Problems with homogeneous unit digits had the unit digit of both operands smaller (e.g., 21 x 63) or larger (e.g., 48 x 59) than 5. In problems with heterogeneous unit digits, the unit digit of one operand was smaller than 5 while that of the other operand was larger than 5 (e.g., 32 x 69). These two types of problems were included because previous studies found that all participants – whatever their age – select the best rounding strategy more easily when the size of the unit digits is homogenous than when it is heterogeneous (e.g., LeFevre, Greenham, & Waheed, 1993). Moreover, half the problems with either homogeneous or heterogeneous unit digits were so-called rounding-down problems because they were best estimated (i.e., closest products from the correct products) with the rounding-down strategy (e.g., 86 x 21) and half rounding-up problems because they were best estimated with the rounding-up strategy (e.g., 64 x 49). Homogeneous and heterogeneous problems had comparable exact products when solved with each rounding strategy (*means* = 3372 and 3342 for homogeneous and heterogeneous problems, respectively).

Finally, given some effects that are known in arithmetic (see Kadosh & Dowker, 2015, for an overview), the following factors were controlled: (a) no operand had a 0 or a 5 unit digit, (b) digits were not repeated within operands (e.g., 22 x 54), (c) digits were not repeated in the same unit or decade positions across operands (e.g., 52 x 57), (d) no reverse order of operands was used (e.g., 46 x 23 vs. 23 x 46), (e) no tie problems such as 24 x 24, were used, (f) the first operand was larger than the second operand in half the problems, and (g) the operand with the smallest unit digit was in the left position in half the problems (e.g., 42 x 36) and in the right position in the other problems (e.g., 23 x 41).

Procedure

Institutional Review Board approval was obtained from the local ethic committee before data collection began (protocol number: 1516-21). Written consents were obtained from participants before the study. Participants were tested individually in a quiet room using a laptop computer equipped with E-prime software (Schneider, Eschman, & Zuccolotto, 2002). They underwent an approximatively 45- to 60-minutes session during which they completed three tasks. These three tasks were administered in the following order: (a) selection-execution task, (b) selection-only task, and (c) execution-only task. These three tasks were administered in that specific order so that the higher cognitive demand of the selection-execution task was further enhanced by the novelty of the task while the lower cognitive demand of the selection-only task was further alleviated by a higher degree of familiarity with the task and with the stimuli (Norman & Shallice, 1986; Stuss et al., 1995). Thereby, we hoped to maximize effects of task demands on participants' metacognitive accuracy, if such effects exist. Also, we wanted to avoid strategy choices in the selection-execution task to be contaminated by strategy execution or strategy selection only, which would have happened if selection-execution task was taken second or third. Finally, in most studies on strategies, participants are usually asked to select and execute strategies. Very rarely are they asked to only select strategies. By asking participants to first select and execute strategies, the present findings on strategy selection could be compared to previous findings so as to determine whether collecting metacognitive judgments results in different findings (e.g., it was important to replicate the findings that strategy selection is influenced by problem features).

Before the computational estimation task, participants were told that their task was to give an approximate answer to each arithmetic problem that is as close as possible to the correct answer without actually calculating the correct answer. To this end, they had to select the best strategy between two available strategies, rounding both operands down (rounding-down strategy) or rounding both operands up (rounding-up strategy) to their closest decades on each problem. The better strategy for a given problem was the strategy that yielded the answer that was closest to the correct product for this problem. In each task, all participants were presented the 32 problems in random order. Moreover, participants had an initial practice phase including six arithmetic problems so that participants could get familiarized with the apparatus and the general procedure used in each of the three tasks. After this initial practice, no participants had difficulties with the procedure. The stimuli were presented in a 60-point Courier black font in the center of the computer screen. Each problem was preceded by a blank screen for 500 ms that was followed by a warning signal ("##") displayed for 400 ms.

Selection-Execution task. The arithmetic problems were presented one by one in the center of the computer screen for 3 sec. each. After each arithmetic problem display, participants were instructed to make an ease of better strategy selection (EoS) judgment.

Specifically, for each problem, participants had to indicate out loud how easy it would be to select the better strategy (between the two available strategies, rounding-up or rounding-down) to estimate the correct product on a 4-point scale (from “very easy” to “very hard”). Problems were presented for three seconds, each followed by a blank screen. Participants had an unlimited time to make their judgments. The experimenter pressed a response key (1, 2, 3, or 4) to move to the next problem as soon as possible after the participant’s response. Once EoS judgments were made on every problem, participants were presented with a computational estimation task. Specifically, the problems were displayed again, one by one in the center of the computer screen for an unlimited duration, and participants were required to choose which of the two available (rounding-down or rounding-up) strategies was the better strategy to estimate the problems, and then to execute it out loud. Instructions emphasized the fact that participants should do nothing more than the initial rounding up or down and multiply the rounded operands. Participants were also told that the mixed-rounding strategy (i.e., rounding one operand down and the other up to the closest decades) was not available. Excluding mixed-rounding strategy aimed at increasing difficulty of strategy choices and, thereby, avoiding ceiling effects. The experimenter pressed a response key immediately after participant’s response. Only the first response was recorded. Self-corrections were not allowed. Each response was immediately followed by assessment of Retrospective Confidence Judgment (RCJ) during which participants were instructed to indicate their level of confidence in whether they selected the best strategy on a 4-point scale (from “not at all confident” to “totally confident”). Participants were asked to make their RCJ out loud. The experimenter pressed a response key (1, 2, 3, or 4) to move on to the next problem directly after the participant’s response. Once again, participants had as much time as they wanted to make their judgments. We recorded participants’ EoS and RCJ judgments on each problem, the number of problems for which they selected the better strategy, the number of errors in strategy execution (i.e., coded 1 if the estimate differed from the expected estimate given the selected strategy, and 0 otherwise), and mean estimation times.

Selection-only task. The selection-execution task was immediately followed by the selection-only task. The two tasks were mostly the same, except that participants were not asked to execute (neither out loud nor covertly) the strategy after selecting it in the selection-only task. After an initial EoS judgment phase, problems were presented one by one in the center of the computer screen for an unlimited time. Participants were required to choose which of the two strategies (rounding-down or rounding-up) was the better strategy to yield the best estimate (defined as the estimate that is closest from exact product). Specifically, they were asked to press the "B" key (standing for "Bas" or "Down" in French) on an AZERTY keyboard if they thought that the rounding-down strategy was the better strategy and to press the "H" key (for "Haut" or "Up" in French) if they thought that the rounding-up strategy was the better strategy. As in the previous block, each response was immediately followed by a RCJ assessment. We recorded the EoS and RCJ judgments for each problem, the number of better strategy selections, and mean selection times.

Execution-only task. In the final block, the cue strategy was displayed above each problem (e.g., RD, 26 x 42), and participants were asked to execute this cued strategy out loud (e.g., "20 x 40 = 800"). Once the given strategy was executed, the experimenter pressed a response key to move on to the next problem as soon as possible after participant's response. Only the first response was recorded. Self-corrections were not allowed. No metacognitive judgments were required. We collected the number of strategy execution errors and mean execution times.

Results

First, we examined age-related differences in strategic variations under each of the three selection-execution, selection-only, and execution-only tasks. Then, we tested the accuracy of participants' EoS judgment and RCJ in the selection-execution and in the selection-only tasks, and compared these judgments across these two tasks. Finally, we determined whether participants were able to use their metacognitive judgments on one problem to adjust their strategy selection on the next problem. Given that older adults had better arithmetic fluency than young adults, we also conducted our analyses with this variable as a covariate. We indicate when the inclusion of the covariate changed our results. Unless otherwise noted, differences were significant to at least $p < .05$. Preliminary analyses indicated that no gender effects were significant on any of the dependent variables.

As classical statistics did not allow us to distinguish between insensitive data (i.e., no significant results) and evidence for the null hypothesis (i.e., a real absence of differences between two conditions), we chose to use Bayesian statistics in addition to more classical frequentist analyses (Dienes, 2014). Bayes factors (B) indicate the relative strength of evidence for two theories (e.g., Jarosz & Wiley, 2014). For this reason, with Bayes factors, one does not have to worry about corrections for multiple testings (see Dienes, 2011). These factors allow three types of conclusions: There is strong evidence for the alternative hypothesis (B much greater than 1); there is strong evidence for the null hypothesis (B close to 0); and the evidence is insensitive (B close to 1). Lee and Wagenmakers (2014) proposed the following decision criteria: Bayes factors greater than 3 or less than 1/3 represent substantial evidence against or for the null hypothesis, respectively. Any B value between 1/3 and 3 indicates that more evidence is needed. In the present experiment, Bayesian tests were conducted using the default conservative priors from our software program (JASP, 2014; see van de Schoot & Depaoli, 2014).

Age-Related Differences in Strategic Variations

First, we analyzed age-related differences in strategy selection (in selection-execution and selection-only tasks) as well in strategy execution (in selection-execution and execution-only tasks) as a function of problem types. The results of these analyses are presented as supplemental results (S1). These data replicated previous findings on how young and older adults select and execute strategies. Like in many studies either in the arithmetic domain or in other cognitive domains (see Lemaire, 2016; Uittenhove & Lemaire, 2015, for overviews), participants were crucially influenced by problem features. They selected the better strategy on each problem when this better strategy was the rounding-up strategy and the problems were homogeneous problems. They executed the rounding-down strategy more quickly on homogeneous problems. Moreover, both young and older adults were influenced by problem features (e.g., both were better at selecting and executing the better strategy on homogeneous problems than on heterogeneous problems). Overall, these findings suggest that collecting metacognitive judgments on each problem in the present experiment has not influenced young and older adults' strategy choices and strategy execution.

Age-Related Differences in Metacognitive Accuracy

After replicating previous findings showing the influence of participants', problems' and tasks' characteristics on strategy selection, the main goal of this study was to test the influence of these three factors on the accuracy of participants' metacognitive judgments for strategy selection. To achieve this end, we calculated a measure of metacognitive accuracy – Area under the ROC curve; A'ROC – based on the EoS and the RCJ in both the selection-execution and selection-only tasks (see Table 1). The A'ROC is a non-parametric measure from signal detection theory which plots the concordances (i.e., a higher judgment on correct better strategy selection or a lower judgment on incorrect better strategy selection) against the discordances (i.e., a higher judgment on incorrect better strategy selection or a lower judgment on incorrect better strategy selection) (Kornbrot, 2006). As 4-point scale was used in the present experiment, three ROC points were computed for each participant. The curve that passes through these points is the ROC curve and the area under this curve was used as an index of metacognitive accuracy. An A'_{ROC} of 0.5 indicated no metacognitive discrimination between better or poorer strategy selections. As opposed to other traditional measures of metacognitive accuracy or resolution (i.e., gamma coefficient, phi coefficient), A'_{ROC} is not contaminated by participants' tendency to give a high or a low confidence rating, allowing to separate metacognitive accuracy from metacognitive bias (Fleming & Lau, 2014). However, for a more straightforward comparison with prior works examining metacognition in aging, we also conducted our analyses using a more traditional measure of metacognition (i.e., gamma coefficient). In the present experiment, these two measures were highly related (r = .67 for the Selection-Execution task, and r = .77 for the Selection-Only task). The analyses carried out on this measure led to the same conclusions as the analyses carried out on the A'_{ROC} index. These results are presented as supplemental results (S2).

Insert Table 1 About Here

We first examined whether individuals were able to make above chance EoS judgments and RCJ. Under both the selection-execution and selection-only tasks, results of the t tests comparing A'_{ROC} indexes to chance showed that the metacognitive coefficients were significantly larger than .50 in both young and older adults for the RCJ, but not for the EoS judgments (see Table 2).

Insert Table 2 About Here

Next, we examined whether the accuracy of participants' metacognitive judgments for strategy selection varied as a function of age, tasks, and stimulus characteristics. Mean A'_{ROC} for both EoS judgments and RCJ were analyzed with mixed-design ANOVAs, 2 (Age: young, older) x 2 (Task: Selection-execution, Selection-only) x 2 (Unit Digit: homogeneous, heterogeneous), with age as the only between-participants factor.

Analyses of EoS judgments showed no main or interaction effects both in the selection-only and in the selection-execution tasks (all *ps*>.18, all *Bs* < 1 and > .30). Concerning the accuracy of RCJ (see Figure 1), the Age x Task, $F(1,61) = 5.48$, $MSe = 1.65$, $\eta^2_p = .08$, $B = 3.23$, the Task x Unit Digit, $F(1,61) = 8.43$, $MSe = 1.47$, $\eta^2_p = .12$, $B = 8.61$, and the Age x Task x Unit Digit interactions, $F(1,61) = 6.89$, $MSe = 1.47$, $\eta^2_p = .10$, $B = 4.67$, were significant. Specifically, results of pairwise comparisons showed that young adults had lower metacognitive accuracy than older adults in the selection-only task, but this effect was larger on homogeneous unit digits (.71 vs. .84), $F(1,61) = 5.75$, $MSe = 3.18$, $B = 3.56$, than on heterogeneous unit digits (.55 vs. .60), $F(1,61) = 3.54$, $MSe = 0.95$, $B = 3.05$. Both age groups had comparable levels of metacognitive accuracy when tested under the selection-execution task, $F = 2.78$; $B = 0.89$. Moreover, older participants (.66) showed higher overall metacognitive accuracy than young participants (.61) when assessed with the A'_{ROC} index, $F(1,61) = 3.96$, $MSe = 2.43$, $\eta^2_p = .06$, $B = 1.289$. However, Bayesian analyses suggest that more evidence is needed before drawing definite conclusions about this overall effect. This probably results from the fact that older adults show better metacognitive accuracy only for the selection-only task and not for the strategy-execution task. Participants also had higher metacognitive accuracy for the selection-only task (.67) than for the selection-execution task (.60), $F(1,61) = 13.39$, $MSe = 1.65$, $\eta^2_p = .17$, $B > 10$. They also had higher metacognitive accuracy on problems with homogeneous unit digits (.71) than on problems with heterogeneous unit digits (.57), $F(1,61) = 45.87$, $MSe = 1.87$, $\eta^2_p = .42$, $B > 10$.

Insert Figure 1 About Here

Age-Related Differences in Metacognitive Regulation

To examine the influence of participants' metacognitive judgments on their subsequent strategy selection, we calculated mean confidence scores for each participant (as participants' EoS accuracy did not differ from chance, the analyses were only conducted on RCJ scores). The mean confidence scores for each participant were used as cut-off scores to determine whether a low-confidence or a high-confidence response was given on each item. This procedure enabled us to control for participants' tendency to give a high or a low confidence rating (response bias). Indeed, with this procedure, the classification of a specific judgment made by one participant as reflecting high- or low-confidence depended on the other judgments made by this participant. Low-confidence responses were coded as 0 and high-confidence responses were coded as 1. To determine whether participants use the result of their metacognitive monitoring to improve their subsequent strategy selection, non-parametric *phi* (*φ)* coefficients were computed between RCJ scores (coded 0; 1) and rates of subsequent better strategy selection (coded 0; 1) for each participant. The correlation between participants' confidence judgments on one problem and better strategy selection on the following problem significantly differed from 0 in both the selection-only task (φ_{young} = -.54 and φ_{old} = -.44) and in the selection-execution tasks (φ_{young} = -.31 and φ_{old} = -.49), all *ps* < .01. These negative correlations indicate that the better strategy is more likely to be selected when a low-confidence response was given on the previous problem and less likely to be selected when a high-confidence response was given on the previous problem. Moreover, to ensure that these correlations did not reflect post-error adjustments (i.e., a cognitive control effect reflecting more careful response strategies after errors; e.g., Kerns et al., 2004; Notebaert et al., 2009), the influence of better strategy selection on current problems was partialled out. All correlations remained significant in both the selection-only ($\varphi_{\text{p-young}}$ = -.36, $\varphi_{\text{p-old}}$ = -.33) and the selection-execution tasks ($\varphi_{\text{p-young}}$ = -30, $\varphi_{\text{p-old}}$ = -.36), all *ps* < .001. Differences in the size of correlations between young and older adults were tested using *R-to-Z* comparisons. None of the age differences were significant, all *ps* > .20

Overall, our results indicate that participants were unable to predict the ease with which they would select the better strategy on each problem but were able to estimate on a trial-by-trial basis the level of confidence associated with their strategy selection, as higher levels of retrospective confidence was associated with higher levels of better strategy selection. Moreover, when participants gave a low-confidence response to a problem, they selected the better strategy more often on the next problem, suggesting that participants were able to use the result of their metacognition to regulate their strategy selection behaviors. Regarding effects of age, young and older adults had comparable metacognitive accuracy when testing under the selection-execution task, and older participants had higher metacognitive accuracy than young participants in the selection-only task, particularly for easier homogeneous problems. Moreover, both young and older adults appear to use the results of this metacognitive monitoring to regulate their selection of the better strategy on the next problem, and to do so to the same extent. Finally, it was interesting to note that, although older adults tended to make more accurate retrospective judgments in the selection-only task than in the selection-execution task (young adults had comparable metacognitive accuracy across these two tasks), both young and older adults were relatively consistent in the accuracy of their metacognitive judgments across selection-execution and selection-only tasks. Correlations in RCJ between these two tasks were r_{overall} = .32, r_{young} = .41, r_{old} = .34 (*p*s < .05).

General Discussion

The present study is the first one to document age-related differences in adults' metacognitive processes for strategy selection in the context of arithmetic problem solving. Young and older participants had to select and execute the better strategy (selection-execution task), to select (without executing it) the better strategy (selection-only task), and to execute (without selecting it) a cued strategy (execution-only task) on each problem. Participants were also asked to prospectively estimate the ease with which they would select the better strategy (EoS judgments) and to retrospectively estimate how confident they are in having selected the better strategy (RCJ) on each problem. Our results replicated previous findings on how young and older adults select and execute strategies. They also revealed how metacognitive processes are involved during strategy selection, and age-related differences or similarities therein. These findings have important implications to further our understanding the role of metacognitive monitoring and metacognitive regulation for strategy selection, and aging effects on these processes.

Metacognition and Strategy Selection

The most important and original finding of this study concerns metacognitive judgments during strategy selection in problem solving tasks. Participants made accurate retrospective, but not prospective, judgments. This result is consistent with previous research in the literature on memory or perceptive metacognition showing that prospective monitoring is generally more difficult and less accurate than retrospective monitoring (e.g., Siedlecka et al., 2016). The present findings suggest that this conclusion extends to other cognitive domains, like problem solving. These differences in accuracy of prospective and retrospective judgments could reflect differences in the cues that participants use to make their judgments, such as response fluency for retrospective judgments and recent outcome history for prospective judgments (Fleming et al., 2016; but see, Besken & Mulligan, 2013; Rhodes & Castel, 2008). Alternatively, some authors suggest that metacognitive judgments are more accurate when they refer to the response already given than when they are about a future response because retrospective confidence results from monitoring the entire decision-making process and could be informed by many more sources of evidence than prospective judgments (e.g., Graziano, Parra, & Singman, 2015; Siedlecka et al., 2016). For instance, in the selection-only task, the speed with which the evaluation of the problems features has been completed before selecting an answer could inform participants' retrospective judgments, but this information is not automatically available for prospective judgments. Either way, our results suggest that participants – whatever their age – did not use or detect available cues before completing the computational estimation tasks to predict their future strategy selection, but were able to do so when they had to evaluate their confidence retrospectively. These results are important because determining whether the better strategy was selected on each problem is a necessary condition for people to be able to interrupt strategy mid-execution in order to switch from a poorer to a better strategy after strategy selection, as postulated by some computational models (e.g., Lovett & Schunn, 1999; Siegler & Araya, 2005). For these reasons, further studies should replicate the present findings and further examine the validity of the original metacognitive judgments that were used in the present experiment. For instance, studies could determine whether the lack of sensitivity of EoS judgments for strategy selection was really due to an inability of participants to predict their future strategy selection, to the fact that EoS judgments were not sensitive enough to capture this ability, or to the variations in problems difficulty on a trial-by-trial basis which could have prevented participants to use recent outcome history with the task to make their predictions. As a first step to document this issue, additional analyses were conducted to determine whether absolute EoS judgments varied as a function of problem characteristics. Results indicated that participants – whatever their age or the task – did not give higher EoS scores on homogenous (M = 2.72) than on heterogeneous (M = 2.64) problems, F<1.This pattern suggests that, in the present study, participants did not use the proximity of the unit digit to 0 as a cue to inform their prospective judgments. Future experiments should further investigate the processes underlying the EoS judgments by manipulating other factors that could have influenced its sensitivity (e.g., presentation duration, participants' ability to estimate digit proximity to 0).

The present analyses also gave us the opportunity to examine strategy regulation via analyses of the relations between metacognitive judgments on the current problems and strategy selection on the immediately following problems. To our knowledge, this study is the first to examine whether metacognitive monitoring on one item can help participants to regulate their performance on the following item. Indeed, all previous studies investigating interactions between metacognitive monitoring and regulation processes examined whether lower judgments on one item lead to the implementation of appropriate strategies on the same item (see Dunlosky & Metcalfe, 2009, for an overview). In the present experiment, participants were able to use the result of their metacognitive evaluation on one item to adjust their strategy selection on the following item. Indeed, when people reported low-confidence in their strategy selection, they increased their rates of better strategy selection on the next problems. This may indicate that low-confidence responses act as a warning signal to pay more attention to information that is useful to select the better strategy on the next problem. Interestingly, this suggests that some aspects of strategy selection on processing current items can be influenced by metacognitive processes executed on the preceding items. Future studies may determine whether such sequential influences are specific to arithmetic and strategy selection, or whether they are also found in other cognitive domains.

Metacognition and Strategy Selection during Aging

Regarding how metacognition changes during aging, our data indicate that accuracy of retrospective judgments was comparable in young and older adults when participants had to select and execute the better strategy; it was even higher in older adults when participants had to only select the better strategy (without executing it). Moreover, both young and older adults were able to use the results of their metacognitive evaluation to regulate their subsequent strategy choices to the same extent. These results are consistent with the hypothesis that some metacognitive processes are spared during aging and are inconsistent with the hypothesis of age-related declines in older adults' metacognitive skills (see reviews by Castel et al., 2016; Hertzog, 2016). Recall that Ardiale and Lemaire (2012, 2013) found that older adults were poorer at within-problem strategy switching. That latter result is inconsistent with the present findings of comparable strategy selection and metacognitive accuracy in young and older adults. It is possible that, although older adults are as able as young adults to judge whether they selected the better strategy after selecting and executing it, when they have to do this while they are engaged in strategy execution, as they did in Ardiale and Lemaire's studies, older adults may have fewer processing resources left free to evaluate on-line whether they have been selecting and are executing the better strategy. The present findings of a decrease in metacognitive accuracy for older – but not young – adults' when the task requires more cognitive resources (i.e., selection-execution task) as compared to when the task is less resource-consuming (selection only) is consistent with this hypothesis. Indeed, this suggests that older adults have optimal metacognition when the task demands are low (i.e., for homogeneous problems in the selection-only task), but their ability to efficiently execute metacognitive processes – namely, to detect and weigh information or cues allowing them to accurately evaluate their cognitive performance – seems to decline as task demands increase (i.e., on heterogeneous problems or in the selection-execution task). If this is the case, it may partly explain why inconsistent findings have been reported in the literature regarding aging effects (or lack thereof) on metacognitive monitoring. For instance, this could possibly explain why Bunnell et al. (1999) found age-related decrease in metacognition when judgments were made on a (resource-consuming) working-memory task while Hertzog et al. (2002) did not find such differences when judgments were made on a (potentially easier) cued-recall memory task (see also Halamish, McGillivray, & Castel, 2011). Further investigations should, of course, corroborate this hypothesis. Future studies should also examine whether the detrimental effects of task demands could be alleviated when crucial cues for accurate metacognitive judgments are made more salient (e.g., Thomas, Bulevich, & Dubois, 2011).

Another issue that should probably be addressed in future studies concerns the surprising metacognitive advantage that we found for older adults over young adults in the selection-only task. As this advantage remained significant even when the influence of the arithmetic fluency was controlled, we can rule out the possibility that this effect is due to age differences in arithmetic skills. This advantage in older adults may possibly result from the fact that older adults have fewer misconceptions (e.g., driven by stereotyped beliefs) or more knowledge than young adults about what is an efficient strategy to solve arithmetic tasks and, thus, are in a better position to judge their performance (as opposed to what is traditionally observed in memory tasks). Indeed, several studies focusing on metacognition have shown that the accuracy of metacognitive judgments in one domain is positively influenced by how skilled we are in this specific domain (e.g., Schmitt & Sha, 2009). Interestingly, the implementation of the processes or knowledge that help older adults to outperform young adults appears to be resource-consuming. Indeed, older adults only show a metacognitive advantage when task demands are low.

One of the important limitations of the present work is that stimuli were repeated across tasks while the order of the tasks was not counterbalanced. Although this design was selected on the basis of previous studies showing no order effect and no influence of problems repetition across blocks on participants' performance, further studies should be conducted to ensure that our results can be replicated using other types of designs. Another limitation of this study is that it does not tell us how metacognitive processes influence strategy selection while young and older adults are selecting the better strategy. The present data clearly showed that young and older adults tend to have an accurately higher level of confidence after selecting the better strategy than after selecting the poorer strategy and tend to have an accurately lower level of confidence after selecting the poorer strategy relative to after selecting the better strategy. However, we do not know how they actually monitor their selection in a situation where they have to select the better strategy and the task does not require them to judge whether they have used the better strategy. So, we ignore whether they wonder if they have just selected the better strategy after strategy selection and while executing the just-selected strategy, or whether they wonder if they have just selected the better strategy before starting to execute strategies. Another important line of research would be to investigate which specific mechanisms account for the reduction observed in older adults' metacognitive accuracy when the task demands are high. Future experiments could be conducted to examine whether effects of aging, task difficulty, and their interaction on metacognitive accuracy correlate with changes in some specific attentional or executive processes. Finally, the present conclusions that young and older adults are able to accurately judge whether they have just selected the better strategy to solve a given problem but are unable to determine whether it will be difficult to select the better strategy before strategy selection need to be replicated both in arithmetic and in other cognitive domains. If this conclusion is confirmed, future works will be in a better position to investigate how mechanisms underlying metacognitive judgments occur before, during, and after strategy selection, and how this changes with age during adulthood. By addressing these issues, computational theories of strategy selection may implement these metacognitive mechanisms and specify how they influence quality of strategy selection and age-related changes therein.

References

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the Tribes of Fluency to Form a Metacognitive Nation. *Personality and Social Psychology Review, 13*, 219–235. doi:10.1177/1088868309341564

Ardiale, E., & Lemaire, P. (2012). Within-item strategy switching: An age comparative study in adults. *Psychology and Aging, 27*, 1138–1151. doi:10.1037/a0027772

Ardiale, E., & Lemaire, P. (2013). Effects of execution duration on within-item strategy switching in young and older adults. *Journal of Cognitive Psychology, 25*, 464–472. doi:10.1080/20445911.2013.789854

Baltes, P. B., & Baltes, M. M. (1990). Psychological perspectives on successful aging: The model of selective optimization with compensation. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 1–34). New York, NY, US: Cambridge University Press.

Besken, M., & Mulligan, N. (2013). Easily perceived, easily remembered? Perceptual interference produces a double dissociation between metamemory and memory performance. *Memory & Cognition, 41*, 897–903. doi:10.3758/s13421-013-0307-8

Bunnell, J. K., Baken, D. M., & Richards-Ward, L. A. (1999). The effect of age on metamemory for working memory. *New Zealand Journal of Psychology, 28*, 23–29.

Castel, A., Middlebrooks, C., & McGillivray, S. (2016). Monitoring memory in old age: Impaired, spared, and aware. In J., Dunlosky, S.K. Tauber (Eds.)*, The Oxford handbook of metamemory*, 463–483.

Cohen, J. (1988). The effect size. *Statistical power analysis for the behavioral sciences*, 77–83

Deltour, J. (1993). Echelle de vocabulaire de Mill Hill de JC Raven. *Adaptation française et normes européennes du Mill Hill et du Standard Progressive Matrices de Raven (PM38). Braine-le-Château: Editions l'application des techniques modernes*.

Dienes, Z. (2011). Bayesian Versus Orthodox Statistics: Which Side Are You On? *Perspectives on Psychological Science, 6*, 274–290. doi:10.1177/1745691611406920

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5*. doi:10.3389/fpsyg.2014.00781

Dodson, C. (2017). Aging and Memory. In J. H. Byrne (Ed.), *Learning and Memory: A Comprehensive Reference* (Vol. 2, pp. 403-421): Academy Press.

Dunlosky, J., & Connor, L. T. (1997). Age differences in the allocation of study time account for age differences in memory performance. *Memory & Cognition, 25*, 691–700. doi:10.3758/bf03211311

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. Thousand Oaks, CA: Sage Publications, Inc.

Eakin, D. K., Hertzog, C., & Harris, W. (2014). Age invariance in semantic and episodic metamemory: Both younger and older adults provide accurate feeling-of-knowing for names of faces. *Aging, Neuropsychology, and Cognition, 21*, 27–51. doi:10.1080/13825585.2013.775217

Fleming, S. M., Massoni, S., Gajdos, T., & Vergnaud, J.-C. (2016). Metacognition about the past and future: quantifying common and distinct influences on prospective and retrospective judgments of self-performance. *Neuroscience of Consciousness, 1*. doi:10.1093/nc/niw018

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in human neuroscience, 8*. doi:10.3389/fnhum.2014.00443

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

French, J. W., Ekstrom, R. B., & Price, L. A. (1963). *Kit of reference tests for cognitive factors*: Educational Testing Service.

Galvin, S. J., Podd, J. V., Drga, V., & Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic Bulletin & Review, 10*, 843–876. doi:10.3758/bf03196546

Geurten, M., & Lemaire, P. (2017). Age-Related Differences in Strategic Monitoring During Arithmetic Problem Solving. *Acta Psychologica*, 180, 105–116. doi:10.1016/j.actpsy.2017.09.005

Geurten, M., & Meulemans, T. (2017). The effect of feedback on children's metacognitive judgments: a heuristic account. *Journal of Cognitive Psychology, 29*, 184–201. doi:10.1080/20445911.2016.1229669

Glisky, E. L. (2007). Changes in cognitive function in human aging *Brain aging: Models, methods, and mechanisms* (pp. 3–20). Boca Raton, FL, US: CRC Press.

Graziano, M., Parra, L. C., & Sigman, M. (2014). Neural Correlates of Perceived Confidence in a Partial Report Paradigm. *Journal of Cognitive Neuroscience, 27*, 1090–1103. doi:10.1162/jocn_a_00759

Halamish, V., McGillivray, S., & Castel, A. D. (2011). Monitoring one's own forgetting in younger and older adults. *Psychology and Aging, 26*, 631–635. doi:10.1037/a0022852

Hertzog, C. (2016). Aging and metacognitive control. In J. Dunlosky & S. K. Tauber (Eds.), *Oxford Handbook of Metacognition*: Oxford University Press.

Hertzog, C., & Dunlosky, J. (2011). Metacognition in later adulthood: Spared monitoring can benefit older adults' self-regulation. *Current Directions in Psychological Science, 20*, 167–173. doi:10.1177/0963721411409026

Hertzog, C., Kidder, D. P., Powell-Moman, A., & Dunlosky, J. (2002). Aging and monitoring associative learning: Is monitoring accuracy spared or impaired? *Psychology and Aging, 17*, 209–225. doi:10.1037/0882-7974.17.2.209

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. doi:10.1037/0882-7974.22.3.607

Hines, J. C., Touron, D. R., & Hertzog, C. (2009). Metacognitive influences on study time allocation in an associative recognition task: An analysis of adult age differences. *Psychology and Aging, 24*, 462–475. doi:10.1037/a0014417

Jarosz, A. F., & Wiley, J. (2014). What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors. *Journal Of Problem Solving*, *7*, 2–9. doi:10.7771/1932-6246.1167

Kadosh, R. C., & Dowker, A. (2015). *The Oxford Handbook of Numerical Cognition*: Oxford University Press, USA.

Kelemen, W., Frost, P., & Weaver, C. (2000). Individual differences in metacognition: Evidence against a general metacognitive ability. *Memory & Cognition, 28*, 92–107. doi:10.3758/BF03211579

Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior Cingulate Conflict Monitoring and Adjustments in Control. *Science, 303*, 1023.

Kimball, D. R., Smith, T. A., & Muntean, W. J. (2012). Does delaying judgments of learning really improve the efficacy of study decisions? Not so much. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 923–954. doi:10.1037/a0026936

Kornbrot, D. E. (2006). Signal detection theory, the approach of choice: Model-based and distribution-free measures and evaluation. *Perception & Psychophysics, 68*, 393–414. doi:10.3758/bf03193685

Kuhlmann, B. G., & Touron, D. R. (2011). Older adults' use of metacognitive knowledge in source monitoring: Spared monitoring but impaired control. *Psychology and Aging, 26*, 143–149. doi:10.1037/a0021055

Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

LeFevre, J.-A., Greenham, S. L., & Waheed, N. (1993). The Development of Procedural and Conceptual Knowledge in Computational Estimation. *Cognition and Instruction, 11*, 95–132. doi:10.1207/s1532690xci1102_1

Lemaire, P. (2016). *Cognitive aging: The role of strategies*. London: Routledge.

Lemaire, P., Arnaud, L., & Lecacheur, M. (2004). Adults' Age-Related Differences in Adaptivity of Strategy Choices: Evidence From Computational Estimation. *Psychology and Aging, 19*, 467–481. doi:10.1037/0882-7974.19.3.467

Lovett, M. C., & Anderson, J. R. (1996). History of Success and Current Context in Problem Solving: Combined Influences on Operator Selection. *Cognitive Psychology, 31*, 168–217. doi:10.1006/cogp.1996.0016

Lovett, M. C., & Schunn, C. D. (1999). Task representations, strategy variability, and base-rate neglect. *Journal of Experimental Psychology: General, 128*, 107–130. doi:10.1037/0096-3445.128.2.107

Metcalfe, J., & Kornell, N. (2003). The Dynamics of Learning and Allocation of Study Time to a Region of Proximal Learning. *Journal of Experimental Psychology: General, 132*, 530–542. doi:10.1037/0096-3445.132.4.530

Metcalfe, J., & Kornell, N. (2005). A Region of Proximal Learning model of study time allocation. *Journal of Memory and Language, 52*, 463–477. doi:10.1016/j.jml.2004.12.001

Murayama, K., Sakaki, M., Yan, V. X., & Smith, G. M. (2014). Type I error inflation in the traditional by-participant analysis to metamemory accuracy: A generalized mixed-effects model perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 1287–1306. doi:10.1037/a0036914

Neches, R. (1987). Learning through incremental refinement of procedures. In D. Klahr, P. Langley, & R. Neches (Eds.), *Production system models of learning and development* (pp. 163–219). Cambridge, MA, US: The MIT Press.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51*, 102–116. doi:10.1037/0003-066X.51.2.102

Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the "labor-in-vain effect.". *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 676–686. doi:10.1037/0278-7393.14.4.676

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (pp. 1–18). New York, NY: Plenum.

Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: An orienting account. *Cognition, 111*, 275–279. doi:10.1016/j.cognition.2009.02.002

Pansky, A., Goldsmith, M., Koriat, A., & Pearlman-Avnion, S. (2009). Memory accuracy in old age: Cognitive, metacognitive, and neurocognitive determinants. *European Journal of Cognitive Psychology, 21*, 303–329. doi:10.1080/09541440802281183

Price, J., Hertzog, C., & Dunlosky, J. (2010). Self-regulated learning in younger and older adults: Does aging affect metacognitive control? *Aging, Neuropsychology, and Cognition, 17*, 329–359. doi:10.1080/13825580903287941

Price, J., & Murray, R. G. (2012). The region of proximal learning heuristic and adult age differences in self-regulated learning. *Psychology and Aging, 27*, 1120–1129. doi:10.1037/a0029860

Rieskamp, J., & Otto, P. E. (2006). SSL: A Theory of How People Learn to Select Strategies. *Journal of Experimental Psychology: General, 135*, 207–236. doi:10.1037/0096-3445.135.2.207

Robinson, A. E., Hertzog, C., & Dunlosky, J. (2006). Aging, Encoding Fluency, and Metacognitive Monitoring. *Aging, Neuropsychology, and Cognition, 13*, 458–478. doi:10.1080/13825580600572983

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General, 137*, 615–625. doi:10.1037/a0013684

Sacher, M., Isingrini, M., & Taconnat, L. (2013). Effects of aging and divided attention on episodic feeling-of-knowing accuracy. *Acta Psychologica, 144*, 258–263. doi:10.1016/j.actpsy.2013.07.004

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User's guide*: Psychology Software Incorporated.

Schmitt, M. C., & Sha, S. (2009). The developmental nature of meta-cognition and the relationship between knowledge and control over time. *Journal of Research in Reading, 32*, 254–271. doi:10.1111/j.1467-9817.2008.01388.x

Siedlecka, M., Paulewicz, B., & Wierzchoń, M. (2016). But I Was So Sure! Metacognitive Judgments Are Less Accurate Given Prospectively than Retrospectively. *Frontiers in Psychology, 7*, 218. doi:10.3389/fpsyg.2016.00218

Siegler, R. S. (2007). Cognitive variability. *Developmental Science, 10*, 104–109. doi:10.1111/j.1467-7687.2007.00571.x

Siegler, R. S., & Araya, R. (2005). A computational model of conscious and unconscious strategy discovery. *Advances in child development and behaviour, 33*, 1–44.

Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. J. S. G. S. Halford (Ed.), *Developing cognitive competence: New approaches to process modeling* (pp. 31–76). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 204–221. doi:10.1037/0278-7393.26.1.204

Souchay, C., & Isingrini, M. (2004). Age related differences in metacognitive control: Role of executive functioning. *Brain and Cognition, 56*, 89–99. doi:10.1016/j.bandc.2004.06.002

Stuss, D. T., Shallice, T., Alexander, M. P., & Picton, T. W. (1995). A Multidisciplinary Approach to Anterior Attentional Functionsa. *Annals of the New York Academy of Sciences, 769*, 191–212. doi:10.1111/j.1749-6632.1995.tb38140.x

Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, 25, 1024–1037. doi:10.1037/0278-7393.25.4.1024

Thomas, A. K., Bulevich, J. B., & Dubois, S. J. (2011). Context affects feeling-of-knowing accuracy in younger and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 96–108. doi:10.1037/a0021612

Tullis, J. G., & Benjamin, A. S. (2012). Consequences of restudy choices in younger and older learners. *Psychonomic Bulletin & Review, 19*, 743–749. doi:10.3758/s13423-012-0266-2

Uittenhove, K., & Lemaire, P. (2015). The Effects of Aging on Numerical Cognition. In R. C. Kadosh & A. Dowker (Eds.), *The Oxford Handbook of Numerical Cognition* (pp. 345–366): Oxford University Press.

van de Schoot, R., & Depaoli, S. (2014). Bayesian analyses: Where to start and what to report. European Health Psychologist, 16, 75–84.

von der Linden, N., Löffler, E., & Schneider, W. (2016). Effects of a Short Strategy Training on Metacognitive Monitoring across the Life-span. *Frontline Learning Research, 3*, 37–55. doi:10.14786/flr.v3i4.196

Table 1

Mean Ease of Selection (EoS) Judgments and Mean Retrospective Confidence Judgments (RCJ) in Young and Older Adults for the Selection-Execution and the Selection-Only Task on Rounding-Down or Rounding-up Problems with Homogenous and Heterogeneous Unit Digits.

Task	Better Strategy	Young Adults			Older Adults			**Total**
		Homogeneous Problems	Heterogeneous Problems	Means	Homogeneous Problems	Heterogeneous Problems	Means	**Means**
					EoS			
Selection-Execution	Rounding Down	1.57 (.05)	2.60 (.05)	2.09 (.04)	1.53 (.07)	1.84 (.07)	1.69 (.05)	1.89 (.03)
	Rounding Up	1.84 (.05)	2.54 (.05)	2.21 (.04)	1.43 (.07)	1.76 (.07)	1.60 (.05)	1.91 (.03)
	Total	1.70 (.04)	2.57 (.04)	2.14 (.03)	1.48 (.05)	1.80 (.05)	1.64 (.03)	1.89 (.03)
					RCJ			
	Rounding Down	3.12 (.05)	2.70 (.05)	2.91 (.04)	3.13 (.06)	2.96 (.06)	3.05 (.04)	2.98 (.03)
	Rounding Up	3.08 (.06)	2.71 (.05)	2.90 (.04)	3.15 (.06)	2.94 (.06)	3.04 (.04)	2.97 (.03)
	Total	3.10 (.04)	2.70 (.04)	2.90 (.03)	3.14 (.04)	2.95 (.04)	3.05 (.03)	2.98 (.03)
					EoS			
Selection-Only	Rounding Down	1.60 (.05)	2.58 (.05)	2.10 (.03)	1.62 (.06)	1.91 (.07)	1.77 (.04)	1.94 (.03)
	Rounding Up	1.92 (.04)	2.49 (.05)	2.20 (.04)	1.54 (.06)	1.46 (.05)	1.50 (.04)	1.85 (.03)
	Total	1.76 (.04)	2.54 (.04)	2.15 (.03)	1.58 (.04)	1.68 (.04)	1.63 (.03)	1.89 (.03)
					RCJ			
	Rounding Down	3.47 (.05)	2.70 (.05)	3.09 (.03)	3.11 (.06)	2.84 (.05)	2.98 (.04)	3.04 (.03)
	Rounding Up	3.56 (.05)	2.75 (.05)	3.16 (.03)	3.16 (.07)	2.89 (.05)	3.03 (.04)	3.10 (.03)
	Total	3.52 (.03)	2.72 (.03)	3.12 (.02)	3.14 (.05)	2.86 (.04)	3.01 (.03)	3.07 (.03)

Note. Standard Errors are in parentheses.

Table 2

Means, Standard Errors, and Tests of Metacognitive Accuracy (A'_{ROC}*) for Young and Older Participants in Selection-Execution and Selection-Only Tasks*

	EoS		RCJ	
	M	*t*	*M*	*t*
	Selection-Execution Task			
Young adults (*N*=37)	.47 (.02)	1.69	.60 (.02)	6.23**
Older adults (*N*=29)	.48 (.02)	0.59	.60 (.02)	3.66*
All (*N*=66)	.48 (.02)	1.44	.60 (.02)	6.71**
	Selection-Only Task			
Young adults (*N*=37)	.48 (.01)	1.74	.63 (.02)	6.46**
Older adults (*N*=29)	.48 (.01)	1.09	.72 (.02)	8.52**
All (*N*=66)	.48 (.01)	1.66	.67 (.02)	10.03**

Note. *t* tests are two-tailed comparisons against chance: 0.5 for Area under the Receiver-Operating-Characteristic (ROC) curve (A'_{ROC}). EoS = Ease of Selection; RCJ = Retrospective Confidence Judgments. $^*p < .05$; $^{**}p < .001$

Figure Caption

Figure 1. Metacognitive accuracy (A'ROC) of retrospective confidence judgments for young and older participants as a function of the unit digit (homogeneous vs. heterogeneous) in Selection-Only (top panel) and Selection-Execution (bottom panel) tasks.