Age-Related Differences in Arithmetic Problem-Verification Strategies

Sandrine Duverne¹ and Patrick Lemaire¹,²

¹Centre National de la Recherche Scientifique–Université de Provence, Marseille.
²Institut Universitaire de France, Marseille.

To test age-related differences in split and problem-difficulty effects, adults between the ages of 20 and 80 years (N = 138) performed a simple and a complex inequality verification task (e.g., 6 + 3 < 11, 271 + 182 < 458; true or false?). Split effects in verification tasks (i.e., better performance for large-split than for small-split problems) reflect strategy selection between nonexhaustive verification (e.g., evaluation of plausibility; estimation) and exhaustive verification (e.g., retrieval; calculation). Problem-difficulty effects (i.e., better performance for easy than hard problems) reflect calculation processing. Results showed decreased split effects across age groups, particularly in the complex task. Moreover, problem-difficulty effects did not vary across age groups. Age-related changes were mostly mediated by age-related declines in processing speed.

PEOPLE use a variety of strategies to accomplish cognitive tasks. As a consequence, they select a given strategy for each problem. How they choose among strategies for each problem has been an important question for cognitive psychologists. Previous research has documented a number of factors that affect strategy selection. Most studies have been carried out in young adults and children (see Siegler, 1996, for a review). As a consequence, we ignore how older adults choose among strategies and which factors affect their strategy choices.

The present research was conducted to address these issues in order to understand age-related differences in strategy selection. Before outlining the logic of the present study, we first review previous findings on age-related differences in strategy use.

A strategy can be defined as “a procedure or a set of procedures for achieving a higher level goal or task” (Jordan & Montani, 1997; Lemaire & Reder, 1999, p. 365). Research in young adults has shown that people vary their strategy use in response to problem and strategy characteristics, their own competence, and task instructions (e.g., DeLoache, 1984; Gardner & Rogoff, 1990; Jordan, Levine, & Huttenlocher, 1995; Jordan & Montani, 1997; Klayman, 1985; Lemaire & Lecacheur, 2002a, 2002b; Reder, 1987). Older adults also use several strategies to accomplish cognitive tasks in a variety of cognitive domains (see Kausler, 1991 and Salihouse, 1991, for reviews). Nevertheless, little research has been conducted regarding adult age-related changes in strategy selection (see, however, Allen, Ashcraft, & Weber, 1992; Duverne, Lemaire, & Michel, 2003; Jordan & Montani, 1997; Lemaire & Lecacheur, 2001; Siegler & Lemaire, 1997).

In arithmetic, Geary and his collaborators analyzed verbal reports of younger and older adults in a simple addition production task (Geary & Wiley, 1991). They observed differences in younger and older adults’ strategy repertoire. Both age groups reported using retrieval (i.e., solving 9 + 4 by directly retrieving 13 from memory) and decomposition strategy (i.e., solving 9 + 4 by doing 10 + 4 − 1). Young adults used decomposition strategy on easier problems (7%) more often than older adults (2%) and retrieval less often than older adults (88% vs. 98%; see Geary, French, & Wiley, 1993, and Geary & Lin, 1998, for similar findings in subtraction problems).

Moreover, older adults retrieved arithmetic facts in memory as fast as younger adults (Allen et al., 1992; Allen, Smith, Jerge, & Viress-Collins, 1997; Geary et al., 1993; Geary & Lin, 1998; Geary & Wiley, 1991; Verhaeghen, Kliegl, & Mayr, 1997; see also Charness & Campbell, 1988, and Salihouse & Coon, 1994, for different results); problem-difficulty effects, the usual index of mental calculation (e.g., Campbell & Graham, 1985; Stazyk, Ashcraft, & Hamann, 1982), did not interact with age. Geary and his colleagues (Geary & Lin, 1998; Geary & Wiley, 1991; Geary et al., 1993) suggested that spared calculation processes with aging may stem from older adults’ compensating for age-related decrements in retrieval processes thanks to higher arithmetic skills acquired during elementary school. Age was shown to affect nonretrieval arithmetic processes by means of declines in processing speed (Salihouse & Coon, 1994).

The present research aimed at investigating adults’ age-related changes in strategy selection and execution in general and in arithmetic problem solving in particular. More specifically, we ask whether younger and older adults are equally able to fine tune their strategy use to problem features. We also ask whether older adults are equally able to execute the specific processes involved in each strategy. Finally, we determined whether processing speed mediates these age-related differences. The target tasks of the present studies are arithmetic problem-verification tasks, both because age-related differences in strategies have not been tested in these tasks and because arithmetic literature offers a rich database in younger adults’ verification strategies.

Previous research reported that younger adults use at least two types of strategies to verify arithmetic problems (i.e., exhaustive and nonexhaustive verification strategies) and that they adjust their strategy use to problem and situation characteristics. The exhaustive verification strategy includes all verification processes such as encoding numbers, calculating the correct solution, comparing and making a true–false decision, and responding. Nonexhaustive strategies do not include all verification processes, such as calculating the correct solution. In previous research, young adults were found to be faster when using nonexhaustive verifications than exhaustive verifications.
One empirical effect is investigated in this project, the split effect, which has been accounted for in terms of nonexhaustive and exhaustive verification strategies. Participants are faster in verifying large-split problems, such as $9 \times 2 = 13$, than in verifying small-split problems, like $9 \times 2 = 17$ (e.g., Ashcraft & Battaglia, 1978). In large-split problems, proposed answers are too far from correct answers to be plausible, enabling participants to quickly respond “false” without calculating the correct answer. As suggested by Ashcraft and Stazyk (1981, p. 191), this magnitude evaluation process “may operate during retrieval to terminate ongoing processes in situations in which various ‘rules of arithmetic’ are violated.” In small-split problems, proposed answers are plausible. Therefore, it is necessary to calculate correct answers and compare calculated and proposed answers before making a true–false decision. Thus, small-split problems are solved with an exhaustive verification strategy and large-split problems with a fast, nonexhaustive strategy. Consistent with such a strategy interpretation, previous research reported that split effects interact with problem difficulty: Small-split latencies increased with problem difficulty more than large-split latencies (Ashcraft & Stazyk, 1981; De Rammelaere, Stuyven, & Vandierendonck, 2001; El Yagoubi, Lemaire, & Besson, 2003). Such interactions arise because the exhaustive verification strategy is influenced by problem difficulty, whereas the nonexhaustive verification strategy is not (see also Lemaire & Fayol, 1995; Lemaire & Reder, 1999). Consequently, analyses of problem-difficulty effects on small-split problems help us to further understand mental calculation processes involved in exhaustive verification strategies.

In the present experiment, age-related differences in split and problem-difficulty effects were tested in 138 adults aged between 20 and 81 years. We used a new arithmetic task whose structure allows us to collect data on simple and complex arithmetic problems, namely the inequality-verification task. Participants were asked to make true–false decisions on 128 problems in a simple task (i.e., two single-digit operands) and on 128 problems in a complex task (i.e., two three-digit operands). Half of the problems were small-split problems (e.g., $5 + 4 < 08$; $284 + 173 < 452$), and half were large-split problems (e.g., $6 + 4 < 03$; $254 + 192 < 407$). Both small- and large-split problems included problems of varying difficulty, based on the size of correct sums and on the presence or absence of carryover of unit digits for simple and complex tasks, respectively. Participants had to indicate whether inequalities were true or false.

Our new inequality-verification task has a number of advantages over equality-verification tasks. First, it is possible to test split effects in complex arithmetic problems (see Geary, Widaman, & Little, 1986) because participants have to calculate hundred and decade digits to make a correct decision. In equality-verification tasks, small- and large-split problems can be correctly solved by just looking at whether the unit digits of proposed sums match those of operands (e.g., in a problem such as $284 + 173 = 452$, the sum of unit digits $4 + 3$ differs from the proposed unit digit 2). Another structural difference between both tasks is that inequality-verification tasks never propose answers equal to correct answers, thereby avoiding previously reported priming retrieval of correct answers (e.g., Campbell, 1987).

The present experiment tested the hypothesis that aging results in differences in strategy use, which makes several predictions. The first set of predictions concerned Age × Split interaction effects and suggested several patterns of age-related differences in strategy selection. First, contrary to young adults, older adults are no longer flexible in their strategy use. This predicts equal levels of speed and accuracy for both small- and large-split problems in older adults. Second, like younger adults, older people use the exhaustive verification and nonexhaustive verification strategies on each type of split problems, but much less adaptively. That is, they are not as systematic as younger adults in their Problem × Strategy assignments, which would be seen in split effects of smaller magnitude across age groups. Alternatively, similar findings in younger, middle-aged, and older adults would show that strategy selection mechanisms remain stable with age and are affected by similar variables in each age group. Finally, we looked at whether age-related changes in strategies varied with task complexity (i.e., one- and three-digit operands for simple and complex tasks, respectively).

The second set of predictions concerned Age × Problem Difficulty interactions. In order to derive more sensitive measures of calculation processes, we performed individual regression analyses (Lorch & Myers, 1990), using a problem-difficulty variable as a predictor. Consistent with previous findings (e.g., Allen et al., 1992; Allen et al., 1997; Geary et al., 1993; Geary & Lin, 1998; Geary & Wiley, 1991), we expected that slopes, which capture central calculation processes (e.g., retrieval of arithmetic facts in memory), would not vary across age groups, whereas intercepts, which capture all the other arithmetic processes that do not vary with problem difficulty (e.g., encoding and responding), would increase with age. Such results may suggest that age-related impairments in mental calculation processes are accounted for by declines in processes other than central calculation processes. Alternatively, age-related increase in intercepts and slopes would reveal that all calculation processes are affected by age.

In addition to performance on simple and complex verification tasks, we assessed individuals’ perceptual processing speed by averaging three paper-and-pencil scores. We performed hierarchical regression analyses, first to evaluate age-related variance in strategy selection and calculation processing, and second to evaluate the amount of age-related variance explained by processing speed. A decreased age-related variance after processing speed has been controlled for may reveal the amount of the effects of age mediated by processing speed (see Salthouse, 2000a, 2000b).

**METHODS**

**Participants**

Participants were 138 individuals: 46 young adults with a mean age of 29.4 years (range 20–39), 46 middle-aged adults with a mean age of 49.3 years (range 40–59), and 46 older adults with a mean age of 69.6 years (range 60–81). Younger adults were undergraduate students from the University of Provence (Aix-en-Provence, France); middle-aged and older adults were recruited from the community. All older adults had scores higher than 27 ($M = 29.4$) on the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975);
Stimuli
Each participant verified two series of 128 simple and complex addition problems separated into two blocks. For both simple and complex tasks, problems were made of 64 pairs of operands each presented twice, once during the first block of the experiment with one split, and once during the second block with the other split. In the simple task, problems corresponded to the 64 possible combinations of digits 2–9; in the complex task, problems were composed of three-digit numbers.

Two types of characteristics were factorially manipulated, that is, problem-split and problem-difficulty characteristics. Based on splits between correct and proposed sums, problems were either small- or large-split problems. More specifically, concerning the simple task, small-split problems had a deviation of ±1 or 2 units from correct sums (e.g., 3 + 8 < 12), whereas large-split problems had a deviation of ±7 or 8 units (e.g., 5 + 7 < 19). In the complex task, small-split problems had sums that were 1% away on average from correct sums (e.g., 157 + 176 < 334), whereas proposed sums for large-split inequalities were 9% (e.g., 168 + 196 < 397). In both simple and complex tasks, half the problems were categorized as “easy problems,” and the other half as “hard problems.” In the simple task, easy problems had correct sums smaller than 10, and hard problems had sums equal to or larger than 10 (see Hamann & Ashcraft, 1985; Kirk & Ashcraft, 2001; LeFevre, Sadesky, & Bisanz, 1996). In the complex task, easy problems involved no carry of the units (e.g., 243 + 121 < 331), and hard problems involved a carry of the units (e.g., 218 + 129 < 314; see Geary et al., 1993; Geary & Lin, 1998). For complex inequalities, carryover of tens was not taken into account to reflect problem difficulty because they were expected to be processed on both small- and large-split problems, and, consequently, they were not meaningful to discriminate strategy use on each split problem.

Following previous findings in arithmetic (see Dehaene, 1997; Geary, 1994, for recent overviews), problems were selected in each Problem Difficulty × Split condition with several constraints. In the simple task, (a) half the problems had proposed sums greater than the correct sums; (b) half the problems had their larger operand on the left position (e.g., 8 + 2 < 11); (c) half the problems had proposed sums whose odd-
even status matched those of correct sums; (d) the number of tie problems (i.e., both operands being equal; e.g., 2 + 2 < 06) was equal in each condition; (e) none of the operands were equal to 0 or 1; (f) none of the proposed sums were equal to correct products of the operands; (g) none of the proposed sums were equal to 0; (h) none of the proposed sums were equal to one of the operands; (i) proposed sums were even for half the problems; and (j) all problems had a proposed sum that was composed of two digits (i.e., one-digit sums were preceded by a 0, such as in 2 + 3 < 07).

As for simple inequalities, complex inequalities were selected in order to control (a) the type of (positive vs. negative) inequalities and (b) the (left–right) positions of the larger operand. Moreover, (c) proposed sums were always within the same hundred as correct sums; (d) half the inequalities involved a carry of the tens (e.g., 284 + 173 < 456); (e) none of the operands had the tens or unit digits equal to 0; (f) none of the operands had unit digits equal to 5; (g) none of the pairs of operands had the same unit digits; and (h) small- and large-split inequalities had the exact same mean correct sums (mean sizes = 402).

Procedure
Participants were tested individually in one session that lasted approximately 2 hr. They first performed the inequality-verification task and then a set of tasks assessing their processing speed. In the inequality-verification tasks, half of the participants first performed the simple task, and then the complex task; and half did the reverse order. Each task was preceded by 16 training problems whose structure was identical to those of experimental stimuli. Participants had short rest periods between the two series of inequalities and between the two blocks of each series. Problems were presented in a 48-point Palatino font in the center of a Macintosh computer screen. Each trial was preceded by an intertrial interval of 2,000 ms, and then a fixation point (“+”) was displayed in the center of the screen for 750 ms. The inequalities were then displayed horizontally in the center of the screen in the form of \(a + b < c\). The symbols and numbers were separated by spaces equal to

Thus, none were excluded from the study. At the end of the experiment, participants completed both the addition and the subtraction–multiplication subtests of the French Kit (French, Ekstrom, & Price, 1963), which provided an assessment of participants’ arithmetic skill with an independent, paper-and-pencil test. Each subtest consisted of two pages of problems. All participants were given 2 min/page and were instructed to solve the problems as fast and accurately as possible. The number of correct answers on each of the addition and the subtraction–multiplication tests were summed to yield a total arithmetic score. Next, participants completed the French version of the Mill Hill Vocabulary Scale (MHVS; Deltour, 1993; Raven, 1951) so as to control for their verbal ability. It consists of 33 items distributed across three pages. Each item was a target word followed by six proposed words, and the task consisted in identifying which of the proposed words had the same meaning as the target word. The number of correct items represented the level of verbal ability. Participants’ characteristics are summarized in Table 1.

### Table 1. Participants’ Characteristics

<table>
<thead>
<tr>
<th>Variables</th>
<th>Young Adults</th>
<th>Middle-Aged Adults</th>
<th>Older Adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years, months)</td>
<td>29.4 ± 7.3</td>
<td>49.3 ± 5.4</td>
<td>69.6 ± 6.4</td>
</tr>
<tr>
<td>Years of education</td>
<td>14.5 ± 1.8</td>
<td>14.7 ± 2.3</td>
<td>13.10 ± 2.2</td>
</tr>
<tr>
<td>Health</td>
<td>3.8 ± 0.7</td>
<td>3.5 ± 0.8</td>
<td>3.5 ± 0.7</td>
</tr>
<tr>
<td>MHVS</td>
<td>25.2 ± 3.3</td>
<td>27.5 ± 3.7</td>
<td>28.4 ± 2.4</td>
</tr>
<tr>
<td>Addition and Subtraction–Multiplication Skills</td>
<td>71 ± 8.2</td>
<td>91 ± 8.2</td>
<td>101 ± 8.2</td>
</tr>
<tr>
<td>MMSE</td>
<td>—</td>
<td>—</td>
<td>29.4 ± 0.9</td>
</tr>
<tr>
<td>LCT</td>
<td>26.4 ± 6.8</td>
<td>25.5 ± 6.7</td>
<td>21.4 ± 4.7</td>
</tr>
<tr>
<td>DDST</td>
<td>59.2 ± 9.7</td>
<td>52.4 ± 10.1</td>
<td>42.4 ± 9.1</td>
</tr>
<tr>
<td>DDST</td>
<td>61.1 ± 8.3</td>
<td>56.9 ± 8.3</td>
<td>47.8 ± 10.6</td>
</tr>
<tr>
<td>% female</td>
<td>63 ± 65</td>
<td>72 ± 65</td>
<td>—</td>
</tr>
</tbody>
</table>

Notes: MHVS = Mill Hill Vocabulary Scale; MMSE = Mini-Mental State Examination; LCT = Letter Comparison Task; DDST = Digit–Symbol Substitution Task; DDST = Digit–Digit Substitution Task. *p < .001.
the width of one character. Each equation subtended approximately 1.58° of vertical visual angle for both simple and complex tasks and 7.2° and 12.9° of horizontal visual angle for simple and complex tasks, respectively. Inequalities remained on the screen until participants responded. A clock was started at the onset of the problem and stopped when the participant pressed one of the two buttons of the CMU button box, corresponding to the true or false response. The software (PsyScope) collected data with 1-ms accuracy. The participants were instructed to use their left and right index fingers to respond, and the assignment of response to buttons was counterbalanced across participants. Participants were encouraged to respond as quickly as possible without making mistakes. Neither processing strategies nor problem characteristics were mentioned in the instructions.

In the second part of the experiment, participants’ processing speed capacities were assessed by means of three frequently used tests: the Letter Comparison Task (LCT), the Digit–Symbol Substitution Task (DSST), and the Digit–Digit Substitution Task (DDST). In the LCT, pairs of series including three, six, or nine letters were presented randomly in a column. Participants had to write between the two series “I” if they were strictly identical or “D” if they were different. Participants had 90 s to respond as many series as possible; their scores corresponded to the number of correct items. The DSST (WAIS-R; Wechsler, 1981) consisted of a code composed of nine digits (1–9) that were coupled to a symbol and 100 items (the first 7 being training items) distributed across four lines. Each item was composed of one box that contained a digit and an empty box that had to be completed with the corresponding symbol. The code remained available until the end of testing. Participants had 90 s to complete as many boxes as possible, while respecting the order of items. Participants’ performance was the number of correct items. The DDST was strictly identical to the DSST, except that the code was composed of digits 1–9 paired with other digits 1–9. The individual scores for the LCT, the DSST, and the DDST significantly correlated with age ($r_s$ $> .33$; see Table 1 for mean scores in each age group) and across them ($r_s$ $>.47$).

Note that these three classic paper-and-pencil tasks mainly assess perceptual processing speed (see, however, Piccin & Rabbitt, 1999, for the impact of memory components in coding task performance), which was shown to be one of the main mediators of age-related declines in cognitive performance in general (see MacDonald, Hultsch, Strauss, & Dixon, 2003; Salthouse, 1996) and in arithmetic performance in particular (Salthouse & Coon, 1994).

**RESULTS AND DISCUSSION**

Analyses of the data are reported in three major sections. The first section examines age-related differences in strategy selection as reflected by split effects. The second section analyzes age-related differences in calculation processes as reflected by problem-difficulty effects. The final section investigates the role of processing speed as a mediator of age-related differences in strategy selection and calculation processing.

**Age-Related Differences in Split Effects**

We ran all reported analyses of variance (ANOVA)s on mean latencies of correctly solved problems. We also analyzed medians and z scores in order to control for potentially artifactual interactions (see Faust, Balota, Spieler, & Ferraro, 1999). All three analyses showed approximately similar patterns. Therefore, only analyses of means are reported, and differences between analyses are mentioned. The ANOVA$s$ involved mixed designs, 3 (age: younger, middle-aged, and older adults) $\times$ 2 (problem difficulty: easy or hard problems) $\times$ 2 (split between proposed and correct sums: small-split or large-split problems), with repeated measures on the last two factors, and participants’ arithmetic skill as a covariate. Adjusted mean verification latencies for simple and complex tasks are presented in Figure 1, and adjusted mean percentages of errors are presented in Figure 2.

**Simple arithmetic task.**—The ANOVA$s$ showed main effects of age, $F(2, 134) = 23.74$, MSE = 2381898, and of problem difficulty, $F(1, 134) = 93.52$, MSE = 43000. They also revealed...
a main 213-ms split effect, $F(1, 134) = 51.35$, MSE = 44024, extending split effects to arithmetic inequality verification tasks. The significant Problem Difficulty $\times$ Split interaction, $F(1, 134) = 10.22$, MSE = 27021, showed larger problem-difficulty effects for small-split problems. Split effects in the present inequality-verification tasks can be interpreted in line with split effects previously observed in equality-verification tasks (e.g., Ashcraft & Stazyk, 1981): People retrieved correct sums in memory before making true–false decisions for small-split problems (i.e., exhaustive verification strategy), and they based their decisions on the implausibility of proposed sums for large-split problems (i.e., nonexhaustive verification strategy). Consistent with this interpretation and with previous findings (e.g., De Rammelaere et al., 2001), small-split problems were more influenced than large-split problems by problem difficulty.

As shown in Figure 1, the Age $\times$ Split interaction was also significant, $F(2, 134) = 8.50$, MSE = 44024. Analyses of split effects in each age group revealed significant 307-ms split effects in younger adults, $F(1, 45) = 108.09$, MSE = 51406; 220-ms split effects in middle-aged adults, $F(1, 45) = 75.54$, MSE = 27460; and 112-ms split effects in older adults, $F(1, 45) = 4.46$, MSE = 65906. In other words, all age groups showed significant split effects in the simple task; these effects decreased in magnitude as age increased. Participants used exhaustive and nonexhaustive verification strategies as a function of splits, but, as people age, they used both strategies less systematically with each type of problem. Moreover, younger adults showed 225-ms problem-difficulty effects, $F(1, 45) = 79.52$, MSE = 47338, just as middle-aged adults showed 269-ms problem-difficulty effects, $F(1, 45) = 67.04$, MSE = 45584, and older adults showed 231-ms effects, $F(1, 45) = 22.78$, MSE = 67316. Finally, in younger adults, the significant Problem Difficulty $\times$ Split interaction, $F(1, 45) = 18.52$, MSE = 23567, revealed larger problem-difficulty effects on small-split problems (307 ms), $F(1, 45) = 18.52$, MSE = 23567, than on large-split problems (144 ms), $F(1, 45) = 26.26$, MSE = 31168. The Difficulty $\times$ Split interaction was marginally significant in middle-aged adults, $F(1, 45) = 2.95$, MSE = 22689, $p = .09$, and in older adults, $F(1, 45) = 3.31$, MSE = 36480, $p = .07$.

None of these results on latency were compromised by speed–accuracy trade-offs, because ANOVAs on percentage of errors, presented in Figure 2, revealed no main or interaction effects ($F$s $< 2.14$). Very few errors were produced by the three age groups (mean percentages of errors were 2.7%, 3.7%, and 4.1% in younger, middle-aged, and older adults, respectively) and for both small- and large-split problems (mean percentages of errors were 4% and 3% on small and large splits, respectively).

Complex arithmetic task.—As in the simple task, ANOVA results in the complex task replicated main effects of age, $F(2, 134) = 26.08$, MSE = 21733254, and problem difficulty, $F(1, 134) = 73.63$, MSE = 1028606. They also showed split effects that were marginally significant with means, $F(1, 134) = 3.53$, MSE = 1009126, $p = .06$, and significant with medians, $F(1, 134) = 5.14$, MSE = 1166078, showing that large-split problems were solved 438 ms faster than small-split problems. Moreover, the significant Problem Difficulty $\times$ Split interaction, $F(1, 134) = 10.72$, MSE = 384004, showed larger problem-difficulty effects on small-split than on large-split problems (1,438 ms vs. 887 ms). Hence, split effects were replicated in the complex task and can be interpreted in terms of different strategies being used on different problem types. Small-split problems were solved with an exhaustive verification strategy (i.e., for each pair of hundreds, tens, and units, people calculated sums and compared them with corresponding sums in the proposed answers before making a true or false decision), and large-split problems were solved with a nonexhaustive verification strategy (i.e., people made true or false decisions without calculating each partial sum). Consistent with this interpretation, nonexhaustive verification strategy was less influenced by problem difficulty than exhaustive verification strategy.

Finally, the following interactions were significant: Age $\times$ Difficulty, $F(2, 134) = 3.55$, MSE = 1028606, and Age $\times$ Split, $F(2, 134) = 9.22$, MSE = 1009126. The Age $\times$ Problem Difficulty $\times$ Split interaction was marginally significant with...
scores, $F(2, 134) = 2.39, \text{MSE} = 474928, p = .09$, and $z$ scores, $F(2, 134) = 2.67, \text{MSE} = 0.037, p = .07$, but not significant with means ($F < 1$). These interaction effects were further analyzed with separate ANOVAs conducted on each age group, with problem difficulty and split as within-subject factors. As we can see in Figure 1, younger adults showed significant 726-ms split effects, $F(1, 45) = 32.28, \text{MSE} = 788809; 946-ms problem-difficulty effects, $F(1, 45) = 52.87, \text{MSE} = 1228465$; and Difficulty $\times$ Split interaction, $F(1, 45) = 23.26, \text{MSE} = 313392$. Middle-aged adults had significant 679-ms split effects, $F(1, 45) = 23.65, \text{MSE} = 935332; 1.036-ms problem-difficulty effects, $F(1, 45) = 56.12, \text{MSE} = 803673$; and Difficulty $\times$ Split interaction, $F(1, 45) = 14.37, \text{MSE} = 271944$. In other words, both younger and middle-aged adults showed significant split effects that interacted with problem difficulty. Older adults showed significant 1,505-ms problem-difficulty effects, $F(1, 45) = 51.08, \text{MSE} = 1544504$, but no significant split effects and Problem Difficulty $\times$ Split interaction ($Fs < 1.5$). The variances between and within older adults in their Strategies $\times$ Problems assignments were too large to reveal any differences across problem types in terms of latencies.

As shown in Figure 2, ANOVAs performed on the percentages of errors revealed the following: (a) a main effect of age, $F(2, 134) = 3.51, \text{MSE} = 169.41$, in that younger adults performed the complex task with fewer errors than middle-aged adults (5.2% vs. 7.6%), $F(2, 89) = 5.14, \text{MSE} = 103.87$, who made as many errors as older adults (7.6% vs. 9.1%; $F < 1$); (b) a 2.1% problem-difficulty effect, $F(1, 134) = 5.24, \text{MSE} = 27.72$; and (c) a 5.1% split effect, $F(1, 134) = 31.08, \text{MSE} = 59.75$. As for latencies, the Problem Difficulty $\times$ Split interaction was also significant, $F(1, 134) = 15.68, \text{MSE} = 24.26$. Finally, the Age $\times$ Problem Difficulty interaction, $F(2, 134) = 3.23, \text{MSE} = 27.72$, showed that problem-difficulty effects were significant only in younger adults (3.5%), $F(1, 45) = 26.72, \text{MSE} = 24.48$, and in middle-aged adults (2.4%), $F(1, 45) = 7.52, \text{MSE} = 32.63$, but not in older adults (0.5%; $F < 1$). As we can see in Figure 2, younger and middle-aged adults committed few errors, except when using the exhaustive verification strategies, especially on hard problems. Older adults globally produced more errors so that there were no differences across problem-difficulty conditions.

**Age-Related Differences in Problem-Difficulty Effects**

We specifically determined calculation processes that were or were not impaired with age, and we analyzed problem-difficulty effects in each age group, a usual index of arithmetic proficiency in the simple arithmetic literature (e.g., Campbell & Graham, 1985; Stazyk et al., 1982). We conducted individual regression analyses for each participant, and we used small-split problems’ mean verification times as the criterion variable and problem-difficulty variable as the predictor. The estimated slope is usually taken to assess central calculation processing, such as searching correct sums in memory, and the estimated intercept is taken to assess other processes, such as encoding and responding (e.g., Allen et al., 1992, 1997; Geary & Wiley, 1991; Salthouse & Coon, 1994).

**Simple arithmetic task.**—The product of the operands was chosen as the predictor variable because previous research showed that this is one of the problem features that best captures retrieval processes in simple addition-verification tasks (e.g., Geary et al., 1986; Miller, Perlmutter, & Keating, 1984; Widaman, Geary, Cormier, & Little, 1989). The mean correlation, over participants, of slopes and intercepts was $r = -.17$, suggesting that intercept and slope coefficients reflect different processes.

A comparison of intercepts across age groups, with arithmetic skill entered as a covariate, revealed significant effects of age on intercepts, $F(2, 134) = 22.54, \text{MSE} = 633892$. Intercepts were larger in middle-aged adults (1,824 ms) than in younger adults (1,439 ms), $F(1, 89) = 10.9, \text{MSE} = 237434$, and even larger in older adults (2,619 ms) than in middle-aged adults, $F(1, 89) = 16.6, \text{MSE} = 872358$. A comparison of slopes across age groups, with arithmetic skill entered as a covariate, did not show significant differences ($F = 1.65$; adjusted mean slopes were 8.2 ms, 8.4 ms, 5.1 ms in younger, middle-aged, and older adults, respectively). This set of results suggests that central calculation processes were not impaired with age. Age-related differences in arithmetic processing were accounted for by impairments of other processes such as encoding or responding.

**Complex arithmetic task.**—The sum of ten and unit digits was kept as the difficulty variable that was the most sensitive, independent measure of calculation processes on complex arithmetic: for example, the difficulty of a problem such as $173 + 281 < 457$ was reflected by $7 + 8 + 3 + 1 = 19$. As in the simple arithmetic task, the mean correlation of slopes and intercepts was $r = -.03$, suggesting that both coefficients capture different processes. The same analyses were performed with the sums of hundred, ten, and unit digits as the difficulty variable as the predictor variable. This problem-difficulty variable predicted mean verification times for small-split problems as well as the sums of ten and unit digits ($R^2$s = .90). However, slopes and intercepts derived with such a variable significantly correlated ($r = -.38$), hence most likely assessing the same underlying cognitive processes. Moreover, sums of ten and unit digits have the advantage of simultaneously reflecting the size of both carryover of tens and units, without taking into account the hundred digits that did not involve any carryover and that were almost always of the same size within each problem.

ANOVAs testing age-related differences on intercepts, with arithmetic skill entered as a covariate, revealed an effect of age, $F(2, 134) = 27.17, \text{MSE} = 3007251$: Intercepts increased in middle-aged adults (4,555 ms) compared with younger adults (3,385 ms), $F(1, 89) = 10.29, \text{MSE} = 3102625$, and in older adults (6,277 ms) compared with middle-aged adults, $F(1, 89) = 20.12, \text{MSE} = 3071455$. Slopes did not vary with age (adjusted means were 145, 142, and 166 ms in younger, middle-aged, and older adults, respectively; $F < 1$). As in simple arithmetic, central calculation processes (e.g., retrieval of arithmetic facts) were maintained with age, whereas the other processes (e.g., encoding and responding) were impaired.

**The Role of Processing Speed as a Mediator of Age-Related Differences**

We ran hierarchical regression analyses with indices of processing speed and indices of arithmetic components that were shown to be affected by age: Strategy selection, as seen in decreased split effects with increasing age, and mental calculation processes, as seen in increased intercepts with age. We
computed processing speed indices by averaging the scores of the three paper-and-pencil tasks that mainly assess perceptual processing speed. We conducted the same analyses with processing speed indices by averaging the $z$ scores from the three paper-and-pencil tasks and by performing a principal component analysis on the three scores. The same results were observed whatever the type of processing speed index used. We entered individuals’ arithmetic skill into the model before age so as to statistically control for it.

**Age-related differences in strategy selection.**—Strategy selection is captured in our inequality verification task by split effects. More specifically, our strategy-selection criterion was calculated as \[\text{[(small-split verification times – large-split verification times)/small-split verification times]}\] for each participant. In the simple arithmetic task, the proportion of age-related variance decreased by 70% (from $R^2 = .208$ to $R^2 = .063$) after control of processing speed; however, age had a unique significant effect after such a control, $F(2, 135) = 13.5$. In the complex arithmetic task, as in the simple arithmetic task, unique effects of age decreased by 69% (from $R^2 = .131$ to $R^2 = .04$) after control of processing speed, and unique effects of age on strategy selection were significant after such a control, $F(2, 135) = 6.3$.

**Age-related differences in arithmetic processes.**—Arithmetic processes that do not vary with problem difficulty such as encoding and responding (i.e., assessed with intercepts in individual regressions) were shown to be affected by age. Age-related variance in intercepts decreased after control of processing speed by 81% (from $R^2 = .223$ to $R^2 = .043$) in the simple arithmetic task, and by 80% (from $R^2 = .286$ to $R^2 = .057$) in the complex arithmetic task. Moreover, age had unique effects after such a control in simple arithmetic, $F(2, 135) = 7.69$, as well as in complex arithmetic, $F(2, 135) = 12.32$.

In sum, split effects decreased in magnitude across age groups, suggesting decreased strategy selection with age. Approximately 70% of this age-related variance in strategy selection was mediated by processing speed, on either simple or complex tasks. Consistent with previous findings (e.g., Geary & Lin, 1998; Geary & Wiley, 1991; Geary et al., 1993) and with results showing that older adults outperformed younger adults on classic paper-and-pencil tasks of basic arithmetic (see Table 1), analyses of problem-difficulty effects revealed that age did not impair central calculation processes (i.e., slopes), but it did impair other arithmetic processes such as encoding and responding (i.e., intercepts). Approximately 80% of age-related variance in arithmetic processes were mediated by processing speed (Salthouse & Coon, 1994). Overall, age-related differences in strategy selection may reflect a conservatism on the part of older adults: Older adults may more frequently use exhaustive verification strategies than younger adults because central calculation processes involved in such strategies are maintained with age.

**Conclusions**

Five important conclusions can be drawn from the present findings on age-related differences in arithmetic problem verification. First, simple and complex inequalities are verified with exhaustive and nonexhaustive verification strategies, as revealed by significant split effects and $\text{Split} \times \text{Difficulty}$ interactions. Second, adults of all ages are able to use both strategies and were adaptive in their strategy selection on a problem-by-problem basis. This was seen in significant split effects in the three age groups in the simple task and significant split effects in younger and middle-aged adults in the complex task. Third, strategy selection is impaired with age, in the sense that adults are less systematic in the $\text{Strategies} \times \text{Problems}$ assignments as they age. In the simple task, this was seen in smaller split effects in middle-aged and older adults than in younger adults. Had middle-aged and older adults used exhaustive and nonexhaustive verification strategies on small- and large-split problems, respectively, as often and as systematically as younger adults, they would have had comparable split effects and $\text{Problem Difficulty} \times \text{Split}$ interactions. In the complex task, age-related differences in $\text{Strategies} \times \text{Problems}$ assignments were even magnified as seen in nonsignificant split effects in older adults. Fourth, age does not affect all processes involved in the exhaustive verification strategy. Calculation processes reflected by problem-difficulty effects are maintained with age, whatever the task complexity, but the other processes are affected. Finally, such age-related differences in strategy selection and arithmetic processes are almost totally mediated by age-related differences in processing speed, more specifically in perceptual processing speed. Nevertheless, effects of age were still significant after these processing speed measures were partialled out. This suggests that other factors mediate age-related differences in strategy selection and arithmetic processes. Future studies will unravel what these factors are.

**Acknowledgments**

This research was funded in part by the Centre National de la Recherche Scientifique (the French National Science Foundation), the Cognitique program of the French Ministère de la Recherche, and the Conseil Régional Provence Alpes Côte d’Azur.

We thank Agnès Arlaud and Anne Sophie Renaud for their help in data collection, as well as two anonymous reviewers and Phil Allen for insightful comments on previous drafts of this article.

Address correspondence to Patrick Lemaire, LPC-CNRS & Université de Provence, 3 Place Victor Hugo, Case 66, 13331 Marseille, France. E-mail: lemaire@up.univ-mrs.fr

**References**


