

Age-related differences in sequential modulations of problem-size and rule-violation effects during arithmetic problem verification tasks

Patrick Lemaire^{1,2} · Fleur Brun¹

Published online: 29 October 2015 © Psychonomic Society, Inc. 2015

Abstract Young and older adults were asked to verify true (e.g., $5 \times 61 = 305$) and false ($5 \times 61 = 315$) arithmetic problems. Half the problems were small (e.g., $5 \times 17 = 85$) and half were large problems (e.g., $5 \times 93 = 465$). Half the false problems respected the five rule (i.e., the product of an operand multiplied by five ends with either 5 or 0), and half violated this rule (e.g., $21 \times 5 = 115$ vs. $21 \times 5 = 113$). Both young and older adults showed problem-size effects (i.e., they verified small problems more quickly than large problems) and fiverule violation effects (i.e., they verified problem violating five rule more quickly than problems respecting five rule). Moreover, we found sequential modulations of these problem-size and five-rule effects. Problem-size effects were larger on current problems following large problems than after small problems, and five-rule violation effects were larger after problems violating the five rule than after no-rule violation problems. Finally, sequential modulations of problemsize effects were larger in older adults than in young adults, and there were no age-related differences in sequential modulations of five-rule violation effects. These findings speak to the determiners of arithmetic performance, as to how well arithmetic calculation and non-calculation strategies are executed and selected on current problems depends on strategies used with preceding problems.

Keywords Aging · Number processing · Arithmetic strategies

Patrick Lemaire patrick.lemaire@univ-amu.fr

¹ Aix-Marseille Université and CNRS, Marseille, France

Introduction

The psychology of arithmetic aims at determining how people solve arithmetic problems as simple as 3 + 4 or 8×7 as well as more complex problems like 123 + 879 or 43×87 and what factors affect participants' performance. Two types of tasks have been used to pursue these goals, namely production and verification tasks. In production tasks, participants are given problems (e.g., $8 \times 7 = ?456 + 348 = ?$) and have to find solutions. In verification tasks, participants are given arithmetic equations (e.g., $8 \times 4 = 32$; $23 \times 4 = 93$) and have to say whether these equations are true or false. In both types of tasks, previous research has found that arithmetic performance is influenced by characteristics of problems, situations, and participants. Also important is the type of strategies that participants use. The present study contributes to the cognitive psychology of arithmetic by documenting sequential effects during arithmetic problem solving and by determining whether these effects change with participants' age during adulthood.

Previous studies in arithmetic found that participants' performance is influenced by the type of problems participants have to solve. For example, participants are faster and/or more accurate while solving smaller problems (i.e., problems with smaller operands like 3×4) than while solving larger problems (i.e., problems with larger operands like 7×8 ; see Zbrodoff & Logan, 2005, for an overview). Also, arithmetic performance varies with problem formats and participants' characteristics. Indeed, participants are faster when problems are presented in the auditory format relative to the visual format (e.g., LeFevre, Lei, Smith-chant, & Mullins, 2001), or when participants are young compared to when they are older (see Duverne & Lemaire, 2005, for a review). Moreover, participants of all ages use several strategies (e.g., counting, retrieval, transformation) when they are asked to find solutions to problems as simple as 3×5 (e.g., LeFevre, Bisanz, Daley,

² LPC-CNRS & Aix-Marseille University, Case D, 3 Place Victor Hugo, 13331 Marseille, France

Buffone, & Sadesky, 1996) or more complex problems as 29 + 56 (e.g., Lemaire & Arnaud, 2008). Variations in arithmetic performance as a function of problems' and participants' characteristics have been accounted for by variations in which strategies participants use and/or how they execute the same strategies on different problems (e.g., Campbell & Xue, 2001). Indeed, participants are slower when they use more complex strategies (e.g., they use repeated addition strategies to solve problems like 3×4 , adding 4 three times) and faster when they use simpler strategies (e.g., they retrieve the solution directly from memory). They are also slower when they use the same strategy (e.g., retrieval) but execute it more slowly on some problems (e.g., large problems like 8×9) than on others (e.g., smaller problems like 3×4).

Finally, participants' arithmetic performance is influenced by situation' characteristics. For example, participants are faster and less accurate under speed pressures (e.g., Beilock & DeCaro, 2007; Campbell & Austin, 2002; Lemaire, Arnaud, & Lecacheur, 2004). One situation characteristic that has recently received some attention concerns problem and strategy transitions. Recent studies found that performance on current problems can be influenced by the type of immediately preceding problems. Similarly, strategy performance on current problems can be influenced by the type of strategies used on preceding problems. For example, Schneider and Anderson (2010) asked participants to verify easier (e.g., 31 + 27 = 58) and harder (29 + 34 = 63) problems. The authors found that participants obtained better performance on current problems if these current problems followed easier problems than after harder problems (see also Hinault, Dufau, & Lemaire, 2014; Lemaire & Hinault, 2014; Uittenhove & Lemaire, 2013a). As another example, Lemaire and Lecacheur (2010) gave participants computational estimation tasks in which participants had to provide approximate products to two-digit multiplication problems like 48×73 . On each problem, participants were instructed which strategy to use. Lemaire and Lecacheur found that participants were faster when they were forced to repeat the same strategy on two consecutive trials than when they had to execute two different strategies (see also Ardiale, Hodzik, & Lemaire, 2012; Lemaire & Leclère 2014a, b; Luwel, Schillemans, Onghena, & Verschaffel, 2009, Schillemans, Luwel, Bulté, Onghena, & Verschaffel, 2009). As a final example, Uittenhove and colleagues (Uittenhove & Lemaire, 2012, 2013a; Uittenhove, Poletti, Dufau, & Lemaire, 2013) found that participants were faster when executing calculation strategies on current problems following execution of easier calculation strategies than after harder calculation strategies on immediately preceding problems. Following Schneider and Anderson (2010), Uittenhove and Lemaire (2012) explained these so-called strategy sequential difficulty effects through lesser availability of executive resources after execution of a difficult strategy or after solving a difficult problem. They proposed that difficult strategies or difficult problems temporarily consume central cognitive resources such as executive functions (e.g., inhibition) and/or working memory. Consequently, the temporary depletion of executive resources results in slowing down execution of the next strategy and/or of solving the next problem. This hypothesis was corroborated by their findings of a correlation between working-memory capacities and the magnitude of sequential difficulty effects (Uittenhove & Lemaire, 2013a) and by larger sequential difficulty effects in populations (like patients with Alzheimer's disease) known to have reduced processing resources (Uittenhove & Lemaire, 2013b). What is not known is whether such sequential effects are found in problem verification tasks when participants use not only calculation strategies but also non calculation strategies, and how young and older adults differ in these sequential effects. Given that noncalculation strategies in verification tasks are very easy to select and execute, they may be immune to sequential effects, in particular in young adults who may have enough available executive resources to not be influenced by these sequential effects. We address these issues in the present experiment where young and older participants were given arithmetic problem verification tasks.

In arithmetic problem verification tasks, participants use several strategies, and their performance depends on the strategies they use. For example, Lemaire and Reder (1999; see also Hinault, Dufau, & Lemaire, 2015; Hinault, Tiberghein, & Lemaire, 2015; Krueger, 1986; Lemaire & Fayol, 1995; Lemaire & Siegler, 1995) found that participants are faster when they solve problems that violate five rule (i.e., products of problems including five as an operand end with either five or zero; e.g., $5 \times 14 = 62$) relative to problems that respect five rule (e.g., $14 \times 5 = 60$). This suggests that participants use a fast, five-rule checking strategy on five-rule violation problems. In contrast, when they verify false, no-rule violation problems, participants calculate the correct answer before comparing the calculated and proposed answers to decide whether the equation is true or false. Similarly, when they verify complex true arithmetic problems like $14 \times 5 = 70$ or $16 \times 7 = 112$, participants use calculation strategies. Such calculation strategies lead them to have better performance for verifying smaller, easier problems (e.g., 13×6 = 78) relative to larger, harder problems (e.g., $47 \times 8 = 376$). Thus, because participants use both calculation and noncalculation strategies in arithmetic problem verification tasks, and because we do not know whether participants' performance and strategy use on current problems are influenced by problems and strategies on preceding problems, here we tested sequential effects in young and older adults while they were verifying true and false equations.

The present study was aimed at testing sequential effects during arithmetic problem verification tasks and at comparing changes with adults' age in these sequential effects. We specifically and originally tested sequential effects during true and false problems verification. During true problems verification, we tested whether problem-size effects (i.e., difference in performance between small and large problems) on current true problems varied with size of the preceding problems (i.e., size of problems was defined on the basis of the size of correct products, such that mean correct products ranged from 64 to 265 for small problems and from 285 to 485 for large problems). Following previous works (e.g., Schneider & Anderson, 2010; Uittenhove & Lemaire, 2012; 2013a, b), sequential effects would be seen on true problems if larger problem-size effects were found on current problems following large problems relative to after small problems. Moreover, to test whether sequential effects found on true problems when participants use calculation strategies generalize to false problems on which participants use non-calculation, rule-violation checking strategies, we compared five-rule violation effects on current problems following five-rule violation problems (e.g., $5 \times 14 =$ 62) and after no-rule violation problems (e.g., $5 \times 14 = 60$). Recall that five-rule violation effects are found when participants are faster while verifying false, five-rule violation problems than while verifying false, no-rule violation problems. We tested fiverule violation effects for several reasons. First, they have been much less studied than other strategy-related effects in the context of arithmetic problem verification tasks. However, they can shed important light on what type of rules participants use when they verify arithmetic problems (e.g., participants can check the odd/even status of numbers, the column of unit digits, the splits between proposed and correct answers), and the use of arithmetic rules, like five rule or parity rule, vary with different task parameters. Second, each time five-rule violation effects have been tested, they were large, robust, and systematic in both young and older adults. Third, magnitudes of five-rule violation effects have been found to be influenced by several problem characteristics (e.g., the number of odd/even operands, problem size), exactly like strategy-related effects found either in arithmetic production or verification tasks. Moreover, accounts of fiverule violation effects in terms of strategies have reached consensus in the arithmetic literature, in contrast to some other effects in verification tasks. Thus, they can shed important light on issues of strategies used by participants during arithmetic problem solving when strategies cannot be investigated directly with external behavioral evidence like in problem verification tasks. Sequential modulations of five-rule violation effects would be found if fiverule violation effects are larger after five-rule violation problems than after no-rule violation problems. This is possible if after using the five-rule violation checking strategy on a given problem, participants prepare themselves for using it again on the next problem. If that next problem is a five-rule violation problem, participants can quickly select and execute the five-rule checking strategy, which will increase their verification speed on that problem.

The additional goal of this experiment was to compare changes with adults' age in magnitudes of sequential effects on true and false problems. It was hard to precisely predict whether sequential modulations of problem-size and five-rule violation effects would be different in young and older adults, given results of previous studies. Indeed, Uittenhove and Lemaire (2013a, b) found no age differences in sequential difficulty effects. However, Lemaire and Hinault (2014; see also Ardiale et al., 2012) found that older adults were more influenced than young adults by the strategy used on previous problems when executing a strategy on current problems. Note that none of these previous studies were run in the context of problem verification tasks, which we did here.

Method

Participants

Thirty-four young adults and 34 older adults participated in this experiment. Participants' characteristics are presented in Table 1.

All older adults completed the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) for potential dementia screening. All individuals had scores higher than 27 (mean: 28.9). Therefore, none were excluded from the study. Next, participants completed a home-made arithmetic fluency test that was devised following the French Kit (French, Ekstrom, & Price, 1963) so as to have an independent, pencil-and-paper assessment of their arithmetic fluency. This test included series of addition, subtraction, and multiplication subtests. Each subtest consisted of one page of problems for a total of three pages. All participants were given 2 min per page, and were instructed to solve as many problems as possible. The number of correct answers on each of the subtests was summed to yield a total arithmetic score (maximum = 180). Each subtest was preceded by ten training problems without time constraint. At the end of the experiment, participants completed a French version of the Mill-Hill Vocabulary Scale (MHVS; Deltour, 1993; Raven, 1951) to assess their verbal fluency. The MHVS consists of 33 items distributed across three pages. Each item was a target word followed by six proposed words, and the task consisted of 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.

Stimuli

The experiment included 240 trials, each including two multiplication problems (for a total of 480 problems). Each problem was presented in a standard form ($\underline{a} \times \underline{b} = \underline{c}$) with \underline{a} as a single digit and \underline{b} as a double digit, or reversed. The single digit was 5 whereas double digits ranged from 13 to 97. Half the 480 problems were true problems, and half were false

Table 1 Participants' characteristics

	Young adults	Older adults	Means	F	
N (Females)	34 (21)	34 (23)	_	-	
Mean age (years; months)	20; 0	73; 1	46; 6	1413.8*	
Age range	16;7 - 25;0	64;11 - 96;6	_	_	
Arithmetic fluency (SD)	44.1 (14;4)	45.8 (10.1)	45.0 (12.4)	0.3	
MHVS (SD)	20.6 (3.9)	24.4 (5.0)	22.8 (4.9)	10.1*	

Note. * *p* < .05

problems. Problems were either small or large problems, and false problems either violated the five rule or did not violate the five rule. Correct products were 163 (range: 65—265) on average for small problems and 387 (range: 285—485) on average for large problems.

Half of the false problems respected the five rule, whereas the other half violated the five rule. The five rule states that the product of an operand multiplied by five ends with either 5 or 0. False problems were created by varying splits (i.e., differences in magnitude between correct products and proposed products). The false five-rule violation problems had products with splits of ± 12 or ± 18 (e.g., $97 \times 5 = 473$, $5 \times 27 = 153$), while the false no-rule violation problems had products with splits of ± 10 or ± 20 (e.g., $5 \times 61 = 315$, $19 \times 5 = 75$) between proposed products and correct products. Mean splits were not statistically different for rule-violation and no-rule violations problems. Also, because true problems were the same problems presented with the correct products, there were no differences in problem size between true and false problems, as well as between false rule-violation problems and false no-rule violation problems.

All false problems respected the parity rule which states that when at least one of the two operands is even, the product is even; otherwise, the product is odd. We selected false problems that did not violate the parity rule because previous works showed that when both parity and fives rule are violated on the same problems, participants are faster than when only one of these rules is violated (Hinault, Dufau, & Lemaire, 2015; Hinault, Tiberghien, & Lemaire, 2015). Also, non-five operands were odd operands as previous works found that five-rule violation effects were largest on five × odd-operand problems (e.g., Lemaire & Fayol, 1995; Lemaire & Reder, 1999).

Half the false problems were preceded by true problems, and the other false problems were preceded by false problems. Similarly, half the true problems were preceded by true problems, whereas the other true problems were preceded by false problems. This resulted in 60 trials of each type (i.e., true-true, true-false, false-true, and false-false).

Moreover, to investigate sequential modulations of problem-size effects, there were four types of true-true problems, depending upon the size of first and second problems in each trial. The second small true problems were preceded by small true problems (small/small) in a quarter of true-true problems and by large true problems (large/small) in a quarter of true-true problems. Similarly, the second large true problems were preceded by small true problems (small/large) in a quarter of true-true problems and by large problems (large/ large) in the final quarter of true-true problems.

Finally, to investigate sequential modulations of five-rule violation effects, there were four types of false-false trials, depending upon whether each of the first and second problems in each trial violated or respected the five rule. The second rule-violation problems were preceded by rule- violation problems (rule-violation/rule-violation) in a quarter of false-false trials and by no-rule violation problems (no-rule violation/ rule-violation) in a quarter of false-false trials. Similarly, the second no-rule violation problems were preceded by rule-violation problems (rule-violation) in a quarter of false-false trials. Similarly, the second no-rule violation problems were preceded by rule-violation problems (rule-violation/no-rule violation) in a quarter of false-false problems and by no-rule violation problems (no-rule violation/no-rule violation) in the final quarter of problems.

Based on previous findings in arithmetic, we controlled the following factors (Campbell, 2005; Geary 1994): (a) no double-digit operands had 0 or 5 as unit digits, (b) no double-digit operands had the same unit digit as decade digit (e.g., 33), (c) half the problems had the double digit on the left position (e.g., $27 \times 5 = 135$) and half on the right position (e.g., $5 \times 27 = 135$), (d) size and direction of splits were matched across conditions, (e) there were no repetitions of the same problems within a trial, (f) no false problems had proposed products equal to 100, and (g) size of exact products did not differ across conditions.

Procedure

The problems were presented horizontally in the center of a computer screen in a 48-point Courier New font. Each trial started with a warning-fixation (##) displayed at the center of the screen for 750 ms, followed by the first problem, another warning-fixation, and the second problem. The problems remained on the screen until participants responded. During the inter-trial interval, there was a 1,000-ms blank screen. A clock began timing at the onset of the problem and stopped

when the participant pressed one of two keys on the computer keyboard, corresponding to true or false. The software (E-Prime) controlled stimulus display and latency collection.

Participants were comfortably seated in a quiet room and had to decide whether the problem was true or false when it appeared. They were instructed to press the "S" or the "L" key on an AZERTY keyboard to respond, with both their index fingers, and the assignment of responses to keys was counterbalanced across participants. No particular strategies were mentioned.

Before the experiment, participants were given a block of nine practice (similar but not identical to experimental) trials to familiarize them with the apparatus and procedure. Then they saw four blocks of 60 trials. Participants were individually tested in one session that lasted approximately 60–90 minutes, with a brief rest period between blocks.

Results

Results are reported in two main parts. The first part analyzes performance on the first problems of each trial so as to investigate age-related differences in problem-size and five-rule violation effects. Second, we examined age-related changes in sequential modulations of problem-size and five-rule violation effects.

Age-related differences in problem-size and five-rule violation effects

Problem-size effects Mean correct solution latencies and percentages of errors (see means in Table 2) on first true problems in each trial were analyzed with mixed-design ANOVAs, 2

 Table 2
 Mean correct solution latencies (in ms) and percentages of errors on first problems of each trial for each problem type, in young and older adults

Problems	Younger adults	Older adults	Means
Problem-size effects			
Small	3,665 (3.4)	4,791 (5.2)	4,228 (4.3)
Large	3,760 (3.9)	4,929 (5.9)	4,344 (4.9)
Means	3,712 (3.6)	4,860 (5.6)	4,286 (4.6)
PSE	94 [†] (0.4)	137 [†] (0.7)	116* (0.2)
Five-rule violation effe	ects		
Five-rule violation	1,715 (0.4)	3,017 (1.4)	2,366 (0.9)
No-rule violation	4,366 (7.8)	6,061 (8.9)	5,213 (8.3)
Means	3,040 (4.1)	4,539 (5.1)	3,789 (4.6)
FRVE	2,651** (7.4)	3,044** (5.9)	2,848** (7.5)

Note. PSE problem-size effects: Large problems – Small problems; *FRVE* Five-rule violation effects: No-rule violation problems – Five-rule violation problems

* p < .05; ** p < .01; † p < .10

(Age: young, older adults) × 2 (Problem Size: small, large problems), with repeated measures on the last factor. Young participants were faster than older adults (3,712 ms vs. 4,860 ms; *F*(1, 66) = 17.71, *MS*e = 2527831, η^2_p = .21, *p* < .0001). All participants rejected small true problems faster than large true problems (4,228 ms vs. 4,344 ms; *F*(1,66) = 5.86, *MS*e = 77896, η^2_p = .08, *p* = .018). The Age × Problem Size interaction was not significant, *F*(1,66) = 2.59, *p* = .656. Analyses of mean percentages of errors revealed only a significant main effect of age, *F*(1, 66) = 4.59, *MS*e = 27.2, η^2_p = .06, *p* = .035.

Five-rule violation effects Mean correct solution latencies and percentages of errors on first false problems in each trial were analyzed with mixed-design ANOVAs, 2 (Age: young, older adults) \times 2 (Problem Type: five-rule violation, no-rule violation problems), with repeated measures on the last factor (see means in Table 2). Young adults were faster (3,040 ms) than older adults (4,539 ms), F(1,66) = 34.30, MSe = 225848, $\eta_{p}^{2} = .34, p \le .0001$. All participants rejected no-rule violation problems more slowly than five-rule violation problems (2, 366 ms vs. 5,213 ms; F(1,66) = 436.71, MSe = 631349, $\eta^2_{\rm p}$ = .87, p < .0001). The Age × Problem Type interaction was not significant, F(1,66) = 2.08, MSe = 631349, p = .15. This showed that five-rule violation effects were significant in both age groups (Fs > 205.9) and of equal magnitudes in young (2651 ms) and older adults (3044 ms). Moreover, participants made more errors on no-rule violation problems than on fiverule violation problems (8.3 % vs. 0.9 %; F(1,66) = 68.17, MSe = 27.8, $\eta_p^2 = .51$, p < .0001). Five-rule effects were significant in both age groups (Fs > 33.62) and were of equal magnitudes in young (7.4 %) and in older (7.5 %) adults; the Age \times Problem Type interaction was not significant, F < 1. No other effects came out significant in either solution times or error rates on first problems.

Age-related differences in sequential modulations of problem-size and five-rule violation effects

Sequential modulations of problem-size effects Mean correct solution latencies and percentages of errors on true, second problems in each trial were analyzed with mixed-design ANOVAs, 2 (Age: young, older adults) \times 2 (First Problem Size: small, large problems) \times (Second Problem Size: small, large problems), with repeated measures on the last two factors (see means in Table 3).¹ Older adults were slower than

¹ As suggested by one reviewer, to test for possible inter-trial influences between problems (i.e., effects of second problems in a trial on first problems in the next trial), despite a 1,000-ms blank screen between trials, we ran the same analyses on performance on first problems of a trial as a function of second problems on the previous trial. However, these analyses revealed that both problem-size effects and five-rule violation effects on first problems in a trial were not modulated by second problem in the preceding trial.

Table 3Mean correct solution latencies (in ms) and percentages oferrors on second problems of each trial as a function of problem types,in young and older adults

Second problem	Younger adults	Older adults	Means		
Sequential modulations	of problem-size ef	fects			
First problems: small	problems				
Small	3,701 (5.8)	4,912 (6.2)	4,306 (6.0)		
Large	3,682 (4.9)	4,888 (5.4)	4,285 (5.1)		
Means	3,691 (5.3)	4,900 (5.8)	4,296 (5.6)		
PSE	-19 (-0.8)	-24 (-0.8)	-22 (-0.8)		
First problems: large problems					
Small	3,356 (4.7)	4,597 (5.3)	3,976 (5.0)		
Large	3,721 (4.1)	5,317 (6.5)	4,519 (5.3)		
Means	3,539 (4.4)	4,957 (5.9)	4,248 (5.1)		
PSE	365** (-0.6)	720** (1.3)	543** (0.3)		
Sequential modulations	of five-rule violation	on effects			
First problems: rule-v	violation problems				
Rule-violation	1,339 (0.4)	2,689 (0.4)	2,014 (0.4)		
No-rule violation	4,415 (11.0)	6,249 (8.5)	5,332 (9.7)		
Means	2,877 (5.7)	4,469 (4.4)	3,673 (5.1)		
Frve	3,076 (10.6)	3,559 (8.1)	3,318 (9.4)		
First problems: no-ru	le violation probler	ns			
Rule-violation	1,811 (0.2)	3,136 (1.7)	2,473 (0.9)		
No-rule violation	4,389 (7.9)	6,222 (7.6)	5,306 (7.7)		
Means	3,100 (4.0)	4,679 (4.7)	3,889 (4.3)		
FRVE	2,579 (7.7)	3,086 (5.9)	2,833 (6.8)		
SFRVE	498 (2.9)	473 (2.2)	485 (2.6)		

PSE Problem-size effects: large problems – small problems, *FRVE* fiverule violation effects: no-rule violation problems – five-rule violation problems, *SFRVE* sequential rule-violation effects: five-rule violation effects on current problems after no-rule violation problems on previous problems – five-rule violation effects on current problems after five-rule violation problems on previous problems

* *p* < .05; ** *p* < .01

young adults (4,928 ms vs. 3,615 ms; F(1.66) = 26.17, MSe = 4481664, $\eta_p^2 = .28$, p < .0001), and participants were faster while solving small problems than while solving large problems (4,141 ms vs. 4,402 ms; F(1,66) = 24.15, MSe = 191226; $\eta_{p}^{2} = .27, p < .0001$). Also, most interesting, the First Problem Size \times Second Problem Size interaction was significant, F(1,66) = 33.86, MSe = 159832, η_{p}^{2} = .34, p < .0001. As illustrated in Fig. 1, this interaction revealed sequential modulations of problem-size effects, with non-significant 22-ms problemsize effects when the current problems were preceded by small problems (F < 1) and significant 543-ms problem-size effects when the current problems followed large problems (F(1,66)) = 54.80, MSe = 182721, $\eta_{p}^{2} = .45$, p < .0001). Most interesting, the Age × First Problem Size × Second Problem Size interaction almost reached conventional significance (F(1,66) = 3.44, MSe = 159831, η_p^2 = .05, p = .068). Problemsize effects were non-significant after small problems both in young (-19 ms; F < 1) and in older (-24 ms; F < 1) adults. However, although problem-size effects were significant in both age groups after large problems (see Fig. 1), they were larger in older adults (720 ms; F(1,33) = 36.47, MSe =241785, $\eta^2_{\rm p} = .52$, p < .0001) than in young adults (365 ms; F(1,33) = 18.33, MSe = 123657, $\eta^2_{\rm p} = .36$, $\underline{p} = .00015$). The Group × Problem Size interaction came out significant when ANOVA was restricted to current problems following large problems, F(1,66) = 5.86, MSe = 182721, $\eta^2_{\rm p} = .08$, $\underline{p} =$.018 No main or interaction effects came out significant on mean percent errors (Fs < 1.0).

Sequential modulations of five-rule violation effects Mean correct solution latencies and percentages of errors on false, second problems in each trial were analyzed with mixed-design ANOVAs, 2 (Age: young, older adults) \times 2 (First Problem Type: five-rule violation, no-rule violation problems) \times 2 (Second Problem Type: five-rule violation, no-rule violation problems), with repeated measures on the last two factors (see means in Table 3).

Mean correct solution latencies revealed main effects of age (F(1,66) = 40.05, *MSe* = 4268065, η_p^2 = .38, *p* < .0001), of first problem type (F(1,66) = 10.29, MSe = 309813, $\eta^2_p = .13$, p =.002064), and of second problem type (F(1,66) = 392.48, MSe = 1638478, $\eta_p^2 = .86$, p < .0001). Young adults were faster than older adults (2,988 ms vs. 4,574 ms). All participants were faster on five-rule violation problems (2,244 ms) than on no-rule violation problems (5,319 ms); and participants were faster after five-rule violation problems (3,673 ms) than after no-rule violation problems (3,889 ms). Moreover, the First Problem Type \times Second Problem Type interaction (F(1,66) = 13.24, MSe = 302458, $\eta_{p}^{2} = .17$, p = .00053) revealed significant sequential modulations of five-rule violation effects, such that five-rule violation effects were 485-ms larger after five-rule violation problems than after no-rule violation problems. Finally, and most importantly, the Age × First Problem Type × Second Problem Type interaction was not significant, F < 1. As illustrated in Fig. 2, sequential modulations of five-rule violation effects were comparable in young (498 ms, F(1,33) = 12.97, MSe = 162197, $\eta^2_{\rm p} = .28, p = .00010$) and older (473 ms, F(1,33) = 4.30, MSe =442719, $\eta^2_{p} = .12, p = .0461$) adults.

Analyses of error rates revealed only a significant main effect of second problem type, F(1,66) = 67.51, MSe = 65.6, $\eta^2_p = .51$, p < .0001, as both young and older participants made fewer errors on current problems violating the five rule (0.7 %) than on no-rule violation problems (8.7 %). No other effects came out significant.

General discussion

In arithmetic, research aims at determining factors that influence participants' performance and underlying processes.



Fig. 1 Changes in problem-size effects (i.e., latencies for large problems – latencies for small problems) on current problems following small versus large problems on the preceding problems of each trial

Previous research established that arithmetic performance is influenced by characteristics of participants, problems, and situations. Previous research also found that participants use different strategies, and that strategy use and strategy performance are influenced by these participant, problem, and situation characteristics. The present study replicated two sets of findings and showed new results. We replicated problem-size effects (i.e., poorer performance on large than on small problems) and five-rule violation effects (i.e., better performance on five-rule violation problems than on no-rule violation problems). We originally found sequential modulations of these problem-size and five-rule violation effects. That is, problem-size effects on current problems were larger after large problems and five-rule violation effects were larger after five-rule violation problems. Interestingly, these effects occurred in both young and older adults. Moreover, sequential modulations of problem-size effects were larger in older than in young adults, but there were no age differences in sequential modulations of five-rule violation effects. These findings have important implications for furthering our understanding of the determiners of arithmetic performance, as well as of how arithmetic calculation and non-calculation strategies are executed and selected on current problems in young and older adults.

As in previous studies of arithmetic (see Zbrodoff & Logan, 2005, for a review), here participants obtained poorer performance on large problems than on small problems because it is easier to calculate correct products for small problems like $5 \times 17 = 85$ than for large problems like $5 \times 93 = 465$. Solving large problems involves retrieving arithmetic facts for problems with larger operands that are known to be less easily accessible in long-term memory (see Zbrodoff & Logan, 2005). Similarly, as in previous studies (e.g., Hinault, Dufau, & Lemaire, 2014; Hinault, Tiberghien, & Lemaire, 2015; Lemaire & Reder, 1999; Masse & Lemaire, 2001), here participants were faster at rejecting false equations when the proposed answer violated the five rule (e.g., $5 \times 27 = 113$) than when the proposed answer did not violate the five rule (e.g., 5



Fig. 2 Five-rule violation effects (i.e., latencies for no-rule violation problems – latencies for five-rule violation problems) on current problems following five-rule violation problems versus no-rule violation problems on the preceding problems of each trial in young and older adults

 $\times 32 = 170$). Presumably, on five-rule violation problems, participants used the five-rule violation checking strategy. In this strategy, while encoding problems, participants quickly detect that the proposed answer cannot be true because it violates the five rule; this enables them to quickly say "false" without having to calculate the correct product. In contrast, on no-rule violation problems, participants had to calculate the correct answers before comparing the calculated and proposed answers to decide whether equations are true or false. Thus, these data show that both age groups used calculation strategies on true and false, no-rule violation problems, and solve the five-rule violation problems with the five-rule violation checking strategy. These were pre-requisite conditions to investigate age-related differences in sequential modulations of problem-size and rule-violation effects.²

Sequential modulations of problem-size and five-rule violation effects were the most interesting and original findings in the present experiment. Thus, problem-size effects on current problems were significant only when current problems followed large problems. It was surprising that problem-size effects were not very large on first problems and nonsignificant on second problems when the problems followed small problems. Smaller problem-size effects have often been found in the arithmetic literature when participants solve five problems (e.g., Ashcraft & Battaglia, 1978; Campbell & Graham, 1985). This is probably why problem-size effects in this experiment (in which participants solved only five true and false problems) were smaller than in previous experiments where participants were asked to solve many non-five and a few five problems. Most interesting, problem-size effects here were modulated by the size of the preceding problems. Had we analyzed problem-size effects disregarding the size of the first problems, we would have found the usual problem-size effects (see Zbrodoff & Logan, 2005, for a review), without discovering that problem-size effects on current problems are modulated by the size of the preceding problems. Problemsize effects were significant only when the current problems followed large problems. As found several times in previous studies (e.g., Schneider & Anderson, 2010; Uittenhove & Lemaire, 2012, 2013a, b), this most likely corresponds to sequential difficulty effects. Such sequential difficulty effects can be accounted for by assuming that solving large problems requires more processing resources than solving small problems. This results in fewer available resources to solve next problems. Such temporary depletion of resources results in increased latencies on the next problems. In contrast, when the first problems were small five problems, resource demands did not exceed participants' available resources. As a result, enough resources were available to solve the next small and large problems equally efficiently. Note that it is possible that if previous and current problems were non-five problems, sequential modulations of problem-size effects may also be found. In such a case, because non-five problems are harder than five problems, it may be possible that problem-size effects would be observed on current problems following both small and large problems. However, magnitudes of problemsize effects on current problems should be smaller following small problems than after large problems, a prediction that could be tested in future studies.

Most interesting, sequential modulations of problem-size effects were larger in older than in young adults. Both age groups showed no significant problem-size effects on current problems when current problems followed small problems and large significant problem-size effects after large problems. Problem-size effects on current problems following large problems were larger in older adults (720 ms) than in young adults (365 ms). This can be understood as resulting from decreased processing resources with age during adulthood (e.g., Craik & Salthouse, 2008) that led to larger increased latencies in older adults relative to young adults while solving large problems following large problems. This is the first study that found larger sequential problem-size effects in older adults. Future studies should first try to replicate this finding and further test their reliability (recall that the Age \times First Problem Size × Second Problem Size was marginally significant, p = .068). Future studies will also have to determine whether such findings are specific to the present context of verification tasks where participants had only five problems to verify or whether sequential difficulty effects are also present in other arithmetic problem-solving tasks (and other cognitive tasks).

Sequential modulations were also found for five-rule effects in the present experiment. That is, five-rule effects on current problems were larger when current problems followed five-rule violation problems than after no-rule violation problems. Interestingly, sequential modulations of five-rule effects were of comparable magnitudes in young and older adults. The data showed that sequential modulations of five-rule effects resulted from increased speed on current five-rule violation problems following five-rule violation problems and comparable speed on current no-rule violation problems. Participants likely repeated the five-rule violation checking strategy when two successive problems violated the five rule. Such strategy repetition yielded additional benefits, consistent

² It could be argued that five-rule violation effects found here were the results of our specific heuristics triggered by our set of stimuli. Indeed, all problems with proposed answers that differed from 0 or 5 were false. Therefore, participants may have adopted the following heuristic (i.e., *« if the unit digit of the proposed answer is different from 0 or 5, then say 'false' »*). Note, however, that even if this heuristic may yield five-rule violation effects, it cannot explain the sequential modulations of five-rule violation effects found here. Moreover, the five-rule violation effects found here were of similar magnitudes to those found in previous studies where both five and non-five problems were tested together (e.g., Lemaire & Reder, 1999, Hinault et al., 2015a, b). In these latter experiments, participants could not adopt such a heuristic because they saw true problems with unit digits different from 0 or 5, such that digits 1–9 did not predict «false» answers.

with previous strategy repetition benefits found in production tasks (e.g., Campbell & Arbuthnott, 2010; Lemaire & Lecacheur, 2010; Lemaire & Leclère, 2014a, b; Luwel, Schillemans, Onghena, & Verschaffel, 2009, Schillemans, Luwel, Bulté, Onghena, & Verschaffel, 2009). Such strategy repetition benefits resulted from recent retrieval and execution of the five-rule violation checking strategy influencing subsequent availability for retrieval and execution of that strategy. That is, having just executed the five-rule violation checking strategy on five-rule violation problems made this strategy more available for retrieval and/or execution while solving the next five-rule violation problems, since this strategy was still in a high state of activation. In other words, when participants were executing the five-rule violation checking strategy, if this strategy had just been used, the procedures of this strategy were still activated and, as a consequence, more quickly executed.

Theoretically, our results speak to models of strategy choices (Lovett & Anderson's 1996 ACT-R model; Lovett & Schunn's, 1999 RCCL model; Rieskamp & Otto's, 2006 SSL model; or Siegler & Araya's, 2005 SCADS model). Current computational models share the core assumption that participants use different strategies to accomplish cognitive tasks, arithmetic problem solving included, select strategies on a problem-by-problem basis, and that performance depends on strategies as different strategies yield different levels of speed and accuracy. These models also assume that agerelated changes in strategic behaviors result from age-related changes in relative costs/benefits of available strategies in a given task or domain. The present results are consistent with most of these assumptions (i.e., participants here used different strategies on different problem types, and strategies differed in relative performance). The present results also suggest that computational models of strategies may account for some of the present findings (e.g., sequential modulations of problem-size and five-rule violation effects, age-related differences in some sequential modulations) with a few changes in some assumptions. For example, by assuming that participants try to minimize demands on executive functions to manage costs of changing strategy from one problem to the next, or costs of strategy selection processes in general, current models would account for sequential modulations found here. They would also be able to account for age-related differences and similarities in sequential effects during arithmetic problem solving or during any other cognitive tasks.

References

- Ardiale, E., Hodzik, S., & Lemaire, P. (2012). Aging and strategy switch costs: A study in arithmetic problem solving. L'Année Psychologique, 112, 345-360.
- Ashcraft, M. H., & Battaglia, J. (1978). Evidence for retrieval and decision processes in mental addition. Journal of Experimental Psychology: Human Learning and Memory, 4, 527-538.
- Beilock, S. L., & DeCaro, M. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. Journal of Experimental Psychology: General, 33(6), 983-998.
- Campbell, J. I. D. (2005). Handbook of mathematical cognition. New York: Psychology Press.
- Campbell, J. I. D., & Aruthnott, K. D. (2010). Effects of mixing and cueing simple addition and multiplication. European Journal of Cognitive Psychology, 22(3), 422-442.
- Campbell, J. I. D., & Austin, S. (2002). Effects of response time deadlines on adults' strategy choices for simple addition. Memory & Cognition, 30(6), 988-994.
- Campbell, J. I. D., & Graham, D. J. (1985). Mental multiplication skill: Structure, process, and acquisition. Canadian Journal of Psychology, 39, 338-366.
- Campbell, J. I. D., & Xue, O. (2001). Cognitive arithmetic across cultures. Journal of Experimental Psychology: General, 130, 299-315. doi:10.1037/0096-3445.130.2.299
- Craik, F. I. M., & Salthouse, T. A. (Eds.). (2008). Handbook of Aging and Cognition (3rd ed.). New York: Psychology Press.
- Deltour J. J. (1993). Echelle de vocabulaire de Mill Hill de J.C. 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.
- Duverne, S., & Lemaire, P. (2005). Aging and arithmetic. In J. I. D. Campbell (Ed.), The Handbook of Mathematical Cognition (pp. 397-412). New York: Psychology Press.
- 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(3), 189-198.
- French, J. W., Ekstrom, R. B., & Price, L. A. (1963). Kit of reference tests for cognitive factors. Princeton: Educational Testing Service.
- Geary, D. C. (1994). Children's mathematical development. Washington: American Psychological Association.
- Hinault, T., Dufau, S., & Lemaire, P. (2014). Sequential modulations of poorer-strategy effects during strategy execution: An event-related potential study in arithmetic. Brain and Cognition., 91, 123-130.
- Hinault, T., Dufau, S., & Lemaire, P. (2015a). Strategy combination in human cognition: A behavioral and ERP study in arithmetic. Psychonomic Bulletin & Review, 22(1), 190-199. doi:10.3758/ s13423-014-0656-8
- Hinault, T., Tiberghien, K., & Lemaire, P. (2015b). Age-related differences in plausibility-checking strategies during arithmetic problem verification tasks. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences. doi:10.1093/geronb/ gbu178
- Krueger, L. E. (1986). Why $2 \times 2 = 5$ looks so wrong: On the odd-even rule in product verification. Memory & Cognition, 14(2), 141-149.
- LeFevre, J., Bisanz, J., Daley, K. E., Buffone, L., & Sadesky, G. S. (1996). Multiple routes to solution of single-digit multiplication problems. Journal of Experimental Psychology: General, 125, 284-306.
- LeFevre, J.-A., Lei, Q., Smith-Chant, B. L., & Mullins, D. B. (2001). Multiplication by eye and by ear for Chinese-speaking and Englishspeaking adults. Canadian Journal of Experimental Psychology, 55(4), 277-284.

Compliance with ethical standards

Author note This research was supported by a grant from the Agence Nationale de la Recherche (Grant No. ANR-13-BSH2-0005-01). We would like to thank Meddy Belgacem for his help in collecting the data.

- Lemaire, P., & Arnaud, L. (2008). Lemaire, P., & Arnaud, L. (2008). arithmetic. American Journal of Psychology, 121(1), 1–16.
- Lemaire, P., Arnaud, L., & Lecacheur, M. (2004). Adults' age-related differences in adaptivity of strategy choices: Evidence from computational estimation. *Psychology and Aging*, 10(3), 467–481.
- Lemaire, P., & Fayol, M. (1995). When plausibility judgments supersede fact retrieval: The example of the odd-even effect on product verification. *Memory and Cognition*, 23(1), 34–48.
- Lemaire, P., & Hinault, T. (2014). Age-related differences in sequential modulations of poorer-strategy effects: A study in arithmetic problem solving. *Experimental Psychology*, 61(4), 253–262.
- Lemaire, P., & Lecacheur, M. (2010). Strategy switch costs in arithmetic problem solving. *Memory & Cognition*, 38(3), 322–332.
- Lemaire, P., & Leclère, M. (2014a). Strategy repetition in young and older adults: A study in arithmetic. *Developmental Psychology*, 50(2), 460–468.
- Lemaire, P., & Leclère, M. (2014b). Strategy Selection in Alzheimer patients: A study in arithmetic. *Journal of Experimental and Clinical Neuropsychology*, 36(5), 507–516.
- Lemaire, P., & Reder, L. (1999). What affects strategy selection in arithmetic ? An example of parity and five effects on product verification. *Memory & Cognition*, 22, 364–382.
- Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General*, 124(1), 83–97.
- 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.
- Lovett, M. C., & Schunn, C. D. (1999). Task representations, strategy variability, and base-rate neglect. *Journal of Experimental Psychology: General*, 128, 107–130.
- Luwel, K., Schillemans, V., Onghena, P., & Verschaffel, L. (2009). Does switching between strategies within the same task involve a cost? *British Journal of Psychology*, 100, 753–771.

- Masse, C., & Lemaire, P. (2001). How do people choose among strategies ? A case study of parity and five-rule effects in arithmetical problem solving. *Psychological Research*, 65(1), 28–33.
- Raven, J. C. (1951). *Guide to using progressive matrices. Sets A, Ab, B.* London: Lewis.
- Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135(2), 207–236.
- Schillemans, V., Luwel, K., Bulté, I., Onghena, P., & Verschaffel, L. (2009). The influence of previous strategy use on individuals' subsequent strategy choice: Findings from a numerosity judgement task. *Psychologica Belgica*, 49(4), 191–205.
- Schneider, D. W., & Anderson, J. R. (2010). Asymmetric switch costs as sequential difficulty effects. *Quarterly Journal of Experimental Psychology*, 63, 1873–1894.
- Siegler, R. S., & Arraya, R. (2005). A computational model of conscious and unconscious strategy discovery. In R. V. Kail (Ed.), Advances in child development and behaviour (pp. 1–42). Oxford: Elsevier.
- Uittenhove, K., & Lemaire, P. (2012). Sequential difficulty effects during strategy execution: A study in arithmetic. *Experimental Psychology*, 59(5), 295–301.
- Uittenhove, K., & Lemaire, P. (2013a). Strategy sequential difficulty effects vary with working-memory and response-stimulus intervals: A study in arithmetic. *Acta Psychologica*, 143, 113–118.
- Uittenhove, K., & Lemaire, P. (2013b). Strategy sequential difficulty effects in Alzheimer patients: A study in arithmetic. *Journal of Clinical and Experimental Neuropsychology*, 35(1), 83–89.
- Uittenhove, K., Poletti, C., Dufau, S., & Lemaire, P. (2013). The time course of strategy sequential difficulty effects: An ERP study in arithmetic. *Experimental Brain Research*, 227, 1–8.
- Zbrodoff, N. J., & Logan, G. D. (2005). What everyone finds: The problem size effect. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 331–345). New York: Psychology Press.