

Efficiency of Arithmetic Procedures Modulates the Problem-Size Effect in Subtraction

M. Isabel Núñez-Peña, Angels Colomé and Elisabet Tubau

Universitat de Barcelona (Spain)

Abstract. The aim of this study was to examine whether differences in strategy selection and/or strategy efficiency can explain the modulation of the problem-size effect by arithmetic skill. More specifically, we wondered whether arithmetic skill increases the use of retrieval strategy in large problems, and/or enhances the efficiency of either retrieval or procedural strategies. The performance of highly-skilled (HS) and less highly-skilled (LS) individuals on a subtraction verification task was analyzed according to problem size and to the strategy reported on a trial-by-trial basis after each problem. The problem size effect was larger for LS individuals than for their HS peers, both in response time and in hit rate. Nevertheless, groups did not differ regarding the strategy reported for each subtraction size. As expected, problems in which retrieval strategy was reported were solved more quickly and more accurately than problems solved by procedural strategies. Responses using retrieval strategy were equally fast in the two groups, but HS individuals performed better than LS when using procedural strategies. The results therefore suggest that the differences in behavioral measures between groups might specifically be due to differences in the efficiency of procedural strategies.

Received 21 May 2014; Revised 3 September 2014; Accepted 8 September 2014

Keywords: problem-size effect, arithmetic skill, calculation strategies, verbal reports, subtraction.

For many years, cognitive psychologists have studied the calculation strategies that are used to solve arithmetic problems (Ashcraft & Guillaume, 2009), focusing their attention on the factors that determine strategy selection and efficiency. Strategy selection refers to the selection of one of the available strategies to solve a given problem, while strategy efficiency refers to how fast and accurately this strategy leads to the solution (see Imbo & Vandierendonck, 2008a, b). The efficiency of the different strategies available seems to be the basis of strategy selection (Campbell & Alberts, 2009), which may vary according to the level of arithmetic skill. This study aimed to investigate whether arithmetic skill favors the selection of more efficient calculation strategies or whether it enhances the efficiency of these strategies.

It is generally agreed that individuals rely on two different strategies when performing calculations. The first—referred to as *retrieval*—consists of retrieving the result directly from long-term memory. For example, in the addition $2 + 3$, the result 5 just pops into

your head. The second—referred to as *procedures*—consists of using procedural solutions such as counting or transformation. For example, the problem $7 + 8$ can be solved by counting eight units on from seven, or by transforming it into $8 + 2 + 5$.

In order to study the kind of strategies used, a combination of verbal reports and behavioral measures (Siegler, Adolph, & Lemaire, 1996) is usually applied. Specifically, problems are grouped on the basis of verbal reports and then analyzed using standard behavioral measures. These studies have found that problems in which participants report having used a retrieval strategy are solved faster and with fewer errors than those in which the use of procedures is reported. This has been attributed to the fact that procedures use requires more steps to reach the solution than retrieval strategy. Although the validity of verbal reports as an indicator of the underlying cognitive processing has been questioned (Cooney & Ladd, 1992; Kirk & Ashcraft, 2001; Russo, Johnson, & Stephens, 1989), recent studies both with behavioral (LeFevre, DeStefano, Penner-Wilger, & Daley, 2006; Seyler, Kirk, & Ashcraft, 2003; Smith-Chant & LeFevre, 2003) and psychophysiological measures (Grabner & De Smedt, 2011; Grabner et al., 2009) have shown that, when properly used, they can be a valid source of information. Accordingly, behavioral differences when comparing trials in which the same strategy has been reported may be a signature of strategy efficiency.

More recently, strategy selection has been investigated by means of the operand-recognition paradigm.

Correspondence concerning this article should be addressed to Maria Isabel Núñez-Peña. Department of Behavioral Sciences Methods. Faculty of Psychology. Universitat de Barcelona. Passeig Vall d'Hebron, 171. 08035. Barcelona (Spain). Phone: +34-933125853. Fax: +34-934021359.

E-mail: inunez@ub.edu

This research was supported by the Ministerio Español de Ciencia y Tecnología (grant PSI2009-10364), the Ministerio Español de Economía y Competitividad (grant PSI2012-35703) and the Generalitat de Catalunya (grant SGR2014-177).

This paradigm is based on the assumption that the algorithmic computation performed during the use of procedures impairs the traces in memory of the operands involved in the problem (Thevenot, Barrouillet, & Fayol, 2001; Thevenot, Fanget, & Fayol, 2007). The operands' recognition time after a calculation task and after a number comparison task is compared, since no computation is required in the latter. If procedures have been used, it will be harder to recognize a previously presented number in a calculation task than in a comparison task. But if the calculation task is solved by retrieving the solution from the long-term memory, then operand recognition will be equally easy after comparison and after calculation. Nevertheless, the validity of these inferences has also been challenged, as we discuss below.

Based on these measures, a variety of studies have suggested that strategy selection and strategy efficiency depend on the problem size, the arithmetic operation, and the level of arithmetic ability. The problem-size effect has been widely reported and refers to the fact that response time and errors increase as the size of the operands increases (Groen & Parkman, 1972; Zbrodoff & Logan, 2005). Campbell and Xue (2001) proposed three sources of the problem-size effect related to strategy selection and strategy efficiency. First, small problems (e.g., $2 + 4$) may be solved by retrieval and large problems (e.g., $8 + 6$) may be solved by procedures: hence the increase in response time and error rate found in the latter. Verbal reports have confirmed this explanation, with more frequent retrieval use reported for small than for large problems (Campbell & Xue, 2001). Second, it has been suggested that retrieval may be more efficient — faster and less error-prone — for small than for large problems, so the problem-size effect may be the consequence of differences in the accessibility of the results from long-term memory (Ashcraft & Christy, 1995). This means that retrieval may be used in all problems, but since large problems are less frequent, they are less activated and thus are more difficult to retrieve. And third, procedural strategies could be less efficient — slower and more error-prone — in large than in small problems (LeFevre, Sadesky, & Bisanz, 1996). This means that all problems may be solved by procedures, although these strategies become less efficient when the size of the problem increases. This may be due to the fact that procedures tend to be more complex (or require more steps to reach the solution) for large problems. In short, three alternative explanations have been proposed for the problem size effect, and latency or error rates do not distinguish between them. Hence, other complementary measures must be used. In the experiment reported, we combined behavioral measures with verbal reports.

In terms of the arithmetic operation, it has traditionally been claimed that adults predominantly use retrieval to solve addition and multiplication and procedures to solve subtraction and division (Barrouillet, Mignon, & Thevenot, 2008; Campbell & Xue, 2001; LeFevre et al., 2006; Seyler et al., 2003). Moreover, adults use more retrieval in multiplication than in addition; this is probably because multiplication is learnt by rote at school, which strengthens the association between a problem and its answer in the long-term memory. Whereas addition and multiplication have been investigated in depth, subtraction and division have received less attention among cognitive psychologists. Focusing research on multiplication and addition prevents a comprehensive understanding of the use of strategies, since these well-practiced operations are more frequently solved by retrieval. Therefore, subtractions or divisions are better suited to investigate differences in the efficiency of procedures. To help to solve this gap our study investigated the strategies used and their efficiency specifically during subtraction solving.

With regard to arithmetic ability, it has been suggested that highly skilled individuals (HS) rely more frequently on retrieval than less highly skilled individuals (LS; Imbo, Vandierendonck, & Rosseel, 2007; LeFevre et al., 1996; LeFevre & Kulak, 1994). Indirect support for this proposal can be found in the reduction of the problem-size effect in HS individuals when solving additions and multiplications (Imbo et al., 2007; Núñez-Peña, Gracia-Bafalluy, & Tubau, 2011; Smith-Chant & LeFevre, 2003). Specifically, LS individuals solved small-size problems faster than medium-size ones. The authors interpreted this result as evidence that LS individuals used retrieval only in small-size problems. HS individuals, on the other hand, did not differ between small and medium-sized problems; this finding led the authors to suggest that they used retrieval in both cases (Núñez-Peña et al., 2011). However, the correspondence between behavioral measures and strategy selection is not as clear as initially considered, and some recent studies have concluded that higher skill in arithmetic problem-solving might also lead to the use of more efficient procedures. In a study with expert arithmetical solvers, Fayol and Thevenot (2012) found that presenting the arithmetical sign before the operands (e.g., $+ \text{ followed by } 8 + 6 = ?$) facilitated the solving process compared to a situation in which the whole of the operation was presented at once. This effect was found for additions and subtractions, but not for multiplications; the finding was interpreted as evidence that a procedure was pre-activated by the “+” and “-” signs but not by the “x” sign. The authors concluded that solving additions and subtractions, even small ones, through procedures was much more common in this highly skilled population than previously believed (see also

Barrouillet & Thevenot, 2013 and Campbell, Chen, & Maslany, 2013). High and low skilled participants were not directly compared, although the authors reported that they had not found this effect when the whole population of Psychology students was considered.

In the context of subtractions, to our knowledge only Thevenot, Castel, Fanget, and Fayol (2010) have explored the relationship between arithmetic skill, problem size and strategy selection. In their first experiment, HS and LS arithmetic problem solvers were asked to perform an operand-recognition task after a calculation task and after a comparison task, and were presented with small (e.g., 9–4), medium (e.g., 11–7) and large subtractions (e.g., 41–27). Comparing the performance on the operand-recognition task after the calculation and comparison tasks, the results for small subtractions showed equivalent recognition after both tasks in the two groups, suggesting that both used retrieval. Moreover, performance on the recognition task was worse after calculation than after comparison in large subtractions in both groups, suggesting the use of procedures. Finally, in medium-sized subtractions, recognition was better after comparison than after subtraction in the less-skilled group, but no differences were found in the highly skilled group; the authors interpreted this as evidence of differences in strategy selection by the groups in these subtractions: retrieval for the HS group, and procedures for the LS group. However, Thevenot et al. were unable to find evidence of this differential selection of strategies in the participants' verbal reports (second and third experiments). Individuals, whatever their ability, reported mainly using retrieval in small subtractions, procedures in large subtractions, and a similar percentage of retrieval and procedures in medium-sized subtractions.

The modulation of the problem-size effect may also be due to differences in the efficiency of the strategies. In our view, the discordance between data from verbal reports and those from the operand-recognition paradigm reported by Thevenot et al. (2010) may be attributed to the fact that HS individuals use procedures more efficiently than their less-skilled peers. As a matter of fact, Thevenot et al. compared the subtraction solution times of both skill groups according to the strategies used and found that LS individuals were slower than their HS peers in both retrieval and procedural trials, although no measures of significance were reported. Furthermore, despite the fact that procedures were used in medium-sized problems by a similar percentage in both groups, a detailed look at the kind of procedures used (see their Table 1) shows that the LS group made more use of costly procedures such as counting or decomposition (Baroody, 1983). This choice of procedures may have obliged LS subjects to recall more numbers during the operand-recognition task, and may account for their poorer results. In contrast, HS individuals used more

Table 1. Means and standard errors (in brackets) for response time (in ms) and hit rate (in %) for small and large subtractions in the LS and HS groups

	LS group		HS group	
	Response time	Hit rate	Response time	Hit rate
Small	687 (30)	95 (1.0)	634 (29)	99 (0.3)
Large	847 (54)	82 (2.8)	684 (53)	96 (0.8)

efficient procedural strategies such as transformation or addition reference (Imbo & Vandierendonck, 2008). Notice that these data also corroborate the proposal raised by Fayol and Thevenot (2012) that becoming expert in arithmetic solving implies an ability to use more sophisticated procedures. However, these are all indirect indicators and more investigation is required.

The aim of the present study was to examine whether the differences in the problem-size effect in highly-skilled and less highly-skilled arithmetic problem solvers are due to differences in strategy selection or strategy efficiency. Most previous studies using additions or multiplications attributed the differences between high and low skilled groups to an increase in retrieval use with ability (although see Campbell et al., 2013 for a possible increase in the use of fast procedures in high skilled participants). We wondered whether this would also be the case in subtraction, an operation considered to be more frequently solved through procedures than multiplication or addition. To this end, highly skilled and less highly skilled arithmetic problem solvers were presented with small and large problems in a subtraction verification task and were asked on a trial-by-trial basis which strategy they had used to solve each problem. Based on previous studies on additions and multiplications (Núñez-Peña et al., 2011; Smith-Chant & LeFevre, 2003), we expected that, as the problem size increased, the response time and errors would increase more for LS individuals than for HS.

By also analyzing performance according to participants' self-reports in each problem size and for each skill group, we expected to shed more light on the sources of the performance differences between groups. As suggested above, these differences may be due to three sources: more use of retrieval by HS individuals in large-sized problems, or more efficient use of either retrieval or procedures by this participants' group.

Methods

Participants

Forty-one individuals participated in the experiment. They were selected from a sample of 143 undergraduate students at the University of Barcelona, who

undertook the *Addition and Subtraction Verification Test* and the *Addition Test* from the French Kit (French, Ekstrom, & Price, 1963). A measure of arithmetic ability for each participant was calculated by adding the number of hits in both tests, and two groups were formed according to the individuals' scores. The less highly skilled arithmetic group (LS group) comprised 20 individuals (12 women; age range = 18–27, mean = 20.4, standard deviation = 2.6) who scored below the first quartile in arithmetic ability (score range = 26–41, mean = 35.7, standard deviation = 5.1). The highly skilled arithmetic group (HS group) comprised 21 individuals (13 women; age range = 18–27, mean = 21.3, standard deviation = 2.3) who scored above the third quartile in arithmetic ability (score range = 59–99, mean = 73.1, standard deviation = 12.1). Groups did not differ either in age ($t(39) = 1.085; p = .285$) or gender ($\chi^2(1) = .016; p = .901$). Participants gave written informed consent to participate after the nature of the study had been explained to them and all were paid for their participation.

Materials

The *Addition and Subtraction Verification Test* and the *Addition Test* from the French Kit (French et al., 1963) were used to determine the experimental groups. The *Addition and Subtraction Verification Test* consists of two pages of 60 additions and subtractions involving two two-digit numbers horizontally presented. Subjects were given a 2-minute period to verify as quickly and as accurately as possible the operations presented on each page (e.g., $32 - 11 = 21$; correct/incorrect?). The *Addition Test* consists of two pages of 60 additions involving three numbers of either one or two digits (e.g., $23 + 74 + 8$). Addends are vertically presented. Participants were asked to solve the additions as fast and as accurately as possible during a 2-minute period for each page. The number of problems correctly solved in both tests was added in order to obtain the participants' total scores.

Experimental trials consisted of subtractions in the form $a - b = c$ that were selected from the material provided by Thevenot et al. (2010). We only used their small and medium subtractions, and so in referring to our material, we will refer to them as “small” and “large” respectively. Three of their small subtractions and one medium subtraction were excluded because they were ties in their addition form (e.g., $8 - 4 = 4$). It is generally agreed that subtractions may be solved by reference to the corresponding addition (e.g., $8 - 4 = 4$ can be solved by remembering the tie $4 + 4 = 8$) and ties are solved more quickly and accurately than similar non-tie problems (LeFevre, Shanahan, & DeStefano, 2004). The material retained presented the following

features: (a) small subtractions consisted of subtracting a number between 2 and 6 from a number between 5 and 10; and (b) large subtractions¹ consisted of subtracting a number between 4 and 9 from a number between 11 and 17.

A total of 16 subtractions for each size were used. Each subtraction was presented with the correct and an incorrect solution. Incorrect solutions consisted of adding or subtracting 1 to or from the correct solution, in order to prevent participants verifying the problem by means of the fast plausibility strategy (Duverne & Lemaire, 2005; El Yagoubi, Lemaire, & Besson, 2003, 2005).

Procedure

The experimental session began with a training period consisting of a series of trials similar to those used in the test period. The participant's task was to decide whether the proposed solution to the subtraction was correct or incorrect by pressing the appropriate button on the mouse. Thumbs of both hands were used to give the answer and response buttons were counterbalanced across subjects. The minimum number of trials presented during training was 18 and a learning criterion of 80% of hits was established to ensure participants clearly understood the task. Feedback regarding incorrect responses was given during this period. The training trials were excluded from the analysis. During the test period each participant was presented with 192 subtractions organized into six blocks. Each block comprised 32 subtractions, 16 for each size, half of them with the correct solution and the other half with an incorrect solution. Trials were randomly presented within each block. Participants were given a 1-minute rest between blocks and were asked to press a button to continue. The structure of each trial was as follows: (1) a fixation point was presented for 500 ms; (2) the minuend was presented for 1500 ms; (3) the subtraction sign was presented for 300 ms; (4) the subtrahend was presented for 2500 ms; (5) the proposed solution was presented for 2500 ms or until the response; and (6) a question about strategy use was presented and remained on the screen until the participant gave an answer. The strategy question screen consisted of the words RETRIEVAL and PROCEDURES and participants were instructed to press the left button of the mouse if they had used retrieval and the right button if they had used procedures. An explanation about what retrieval and procedures mean was provided to participants at the beginning of the experiment:

¹Following Seyler, Kirk and Ashcraft (2003), who found a pronounced discontinuity beginning with minuend 11, it is usually considered that this number marks a shift in processing between smaller and larger subtraction facts.

Table 2. Means and standard errors (in brackets) for response time (in ms) and hit rate (in %) for small and large subtractions in the LS and HS groups for correct and incorrect solutions

	Correct solutions				Incorrect solutions			
	LS group		HS group		LS group		HS group	
	Response time	Hit rate	Response time	Hit rate	Response time	Hit rate	Response time	Hit rate
Small	633 (32)	96 (0.7)	597 (31)	99 (0.7)	740 (31)	94 (1.0)	671 (31)	98 (1.0)
Large	792 (56)	83 (1.8)	630 (55)	96 (1.7)	902 (56)	81 (2.5)	737 (54)	96 (2.4)

When you solve a subtraction, you can use different strategies. After each subtraction you will be questioned about which strategy you have used to solve it and you will be given two options: retrieval or procedures. You will have to press the left button if you have used retrieval and you will have to press the right button if you have used procedures. Retrieval consists of just remembering the answer: it just pops into your head. Procedures consist of calculating the answer by using different strategies. For example: (a) you can count down from the first number in the problem ($19 - 3 = 16$ can be solved by counting down from 19: 18, 17, 16); (b) you can count up from the second number until you reach the first number, the answer corresponds to the number of steps ($19 - 16 = 3$ can be solved by counting up from 16: 17, 18, 19); (c) you can decompose the problem ($30 - 13 = 17$ can be solved by decomposing the problem into $30 - 10 - 3$); or (d) you can refer to the corresponding addition ($14 - 6 = 8$ can be solved by remembering the addition $6 + 8 = 14$). These are some examples but they are not all the procedures you can use.

The experiment was controlled by the E-Prime 2.0 program (Psychology Software Tools Inc, Sharpsburg, PA, USA). Numbers were presented in white against a black background, and subtended a visual angle of 1.76° vertically and 1.10° (for one-digit stimuli) or 2.42° (for two-digit stimuli) horizontally.

Results

Solution times and accuracy

Median² response time (RT) for correctly solved trials and the percentage of correct responses were analyzed with analyses of variance (ANOVAs), taking Problem

size (small, large) and Proposed solution (correct, incorrect) as within-subject factors and Group (HS, LS) as the between-subjects factor. The F value, the degrees of freedom, the probability level, and the η_p^2 effect size index (Kirk, 1996) are given. Whenever an interaction reached significance, pairwise comparisons were conducted, and the Bonferroni adjustment was used to control for the increase in Type I error. Only significant effects ($p \leq .05$) are reported.

In terms of response time, the main effects of Problem size ($F(1, 39) = 21.26, p < .001, \eta_p^2 = .35$) and Proposed solution ($F(1, 39) = 53.06, p < .001, \eta_p^2 = .58$) were statistically significant. Small subtractions were verified faster (660 ms) than large subtractions (765 ms) and incorrect problems were solved slower (762 ms) than correct ones (663 ms). The interaction Problem size \times Group was also significant ($F(1, 39) = 5.91, p = .02, \eta_p^2 = .13$). When problem sizes were compared in each group, results showed that both groups were slower in large subtractions than in small subtractions (see Tables 1 and 2). However, analysis of the difference in response time between large and small subtractions in both groups showed that the increase in response time was larger for the LS group (160 ms) than for the HS group (50 ms) ($F(1, 39) = 5.91, p = .02, \eta_p^2 = .13$).

In terms of hit rate, differences were found between groups ($F(1, 39) = 22.7, p < .001, \eta_p^2 = .37$) showing that LS individuals made less hits (89%) than their more skilled counterparts (97%). The main effect of Problem size ($F(1, 39) = 43.1, p < .001, \eta_p^2 = .53$) was also significant, showing that small subtractions were verified more accurately (97% hits) than large subtractions (89% hits). The interaction Problem size \times Group was also significant ($F(1, 39) = 20.1, p < .001, \eta_p^2 = .34$). In order to study this interaction in more detail, problem sizes were compared in each group. Both groups made more errors in large than in small subtractions ($F(1, 19) = 32.28, p < .001, \eta_p^2 = .63$ for the HS group and $F(1, 20) = 11.29, p = .003, \eta_p^2 = .36$ for the LS group; see Tables 1 and 2). However, analysis of the difference in hit rate between small and large subtractions in both groups showed that

²Medians of RT were calculated for each participant and each experimental condition. Medians were used instead of means because the former are more resistant estimators (they are less influenced by outliers).

Table 3. Means and standard errors (in brackets) of percentage of use, response time (in ms) and hit rate (in %) as a function of self-reported strategy, problem size and arithmetic ability

	LS group			HS group		
	% use	Response time	Hit rate	% use	Response time	Hit rate
Direct retrieval						
Small	71 (6.7)	677 (27)	97 (0.7)	69 (7.3)	640 (30)	99 (0.8)
Large	49 (6.7)	796 (56)	85 (2.4)	49 (7.2)	666 (62)	96 (2.6)
Procedures						
Small	29 (6.7)	784 (43)	85 (2.9)	31 (7.3)	660 (47)	98 (3.1)
Large	51 (6.7)	1094 (73)	76 (2.4)	51 (7.3)	777 (79)	93 (2.6)

the drop in hit rate was larger for the LS group (13 %) than for the HS group (2%) ($F(1, 39) = 20.1, p < .001, \eta_p^2 = .34$).

Self-report analysis

The percentage of retrieval use was calculated for each problem size and for each participant, and was analyzed using ANOVA, taking Problem size (small, large) and Proposed solution (correct, incorrect) as within-subject factors and Group (HS, LS) as the between-subjects factor. Data from four HS individuals were not taken into account in this analysis because of technical problems during the recording of their self-report answers. The analysis revealed a significant Problem size effect ($F(1, 35) = 36.5, p < .001, \eta_p^2 = .51$), showing that the percentage of retrieval use was higher in small subtractions than in larger ones (70% vs 49% respectively). Neither the main effect of Proposed solution, nor the main effect of the Group, nor the interactions between them were significant. Crucially, groups did not differ in the percentage of retrieval use for each problem size (71% and 69% in small subtractions for the LS and HS groups respectively; 49% in large subtractions for both LS and HS groups). Table 3 shows the means and standard errors for percentage of strategy use as a function of problem size and arithmetic ability.

Solution times and accuracy as a function of self-reported strategy

We analyzed median response times and hit rates as a function of self-reported strategy, taking Strategy (retrieval, procedures), Problem size (small, large) and Proposed solution (correct, incorrect) as within-subject factors and Group (HS, LS) as the between-subjects factor. Hit rates were calculated by dividing the number of trials correctly answered for a given combination of size and strategy between the number of trials in which the participant reported having used

that strategy³. For example, if a participant reported using retrieval in 25 small subtraction problems and had 23 hits in small subtraction problems in which he reported retrieval use, then the hit rate was 0.92. In the following we discuss only the main effect of strategy and its interaction with other factors, when they reached statistical significance, because the other effects have been reported in the previous analysis.

In terms of response time, the main effect of strategy was significant ($F(1, 35) = 20.10, p < .001, \eta_p^2 = .37$); individuals were slower when they reported having used procedures (829 ms) than when they reported having used retrieval (636 ms), and the interaction Strategy \times Group was also significant ($F(1, 35) = 5.26, p = .028, \eta_p^2 = .13$). Groups did not differ when they reported having used retrieval, but LS individuals were slower (939 ms) than HS individuals (719 ms) when they reported having used procedures ($F(1, 35) = 7.79, p = .008, \eta_p^2 = .18$). The interaction Strategy \times Problem size was also significant ($F(1, 35) = 9.05; p = .005, \eta_p^2 = .21$). The Strategy effect was significant in small ($F(1, 35) = 9.08; p = .005, \eta_p^2 = .21$) and large subtractions ($F(1, 35) = 17.16; p < .001, \eta_p^2 = .33$), showing that retrieval was faster than procedures in both sizes (658.5 ms vs 722 ms in small subtractions; 731 ms vs 935.5 ms in large subtractions). Mean response times and standard errors as a function of self-reported strategy, problem size and arithmetic ability are given in Table 3.

In terms of hit rate, the effect of strategy ($F(1, 35) = 14.97, p < .001, \eta_p^2 = .30$) was significant, showing that procedures were more error-prone (88% of hits) than retrieval (94% of hits). This effect was modulated by the interaction Strategy \times Group ($F(1, 35) = 6.87, p = .013, \eta_p^2 = .16$), which showed that the strategy effect was

³Since not all participants used both retrieval and procedural strategies for all problem sizes, some empty cells occurred in the ANOVAs. Following Imbo and Vandierendonck (2008), these empty cells were replaced with the mean response time or hit rate of the corresponding condition (i.e., the mean response time over participants of the Strategy \times Problem size \times Group corresponding condition per participant).

only significant for LS individuals ($F(1, 19) = 13.8, p = .001, \eta_p^2 = .42$; 91% and 81% for retrieval and procedures respectively) but not for their HS peers (96% and 93% for retrieval and procedures respectively).

Discussion

In the present study we examined the strategies used when solving subtractions by individuals with high- and low-arithmetic ability. Our aim was to investigate whether differences in the problem size effect depending on arithmetic ability are due to differences in strategy selection or strategy efficiency. To this end, individuals with high- and low-arithmetic ability were presented with small and large subtractions and were asked to report the strategy they had used after each problem on a trial-by-trial basis. Behavioral measures were analyzed according to the strategy reported in order to gather further information on the sources of the differences in the problem-size effect in both groups.

Previous studies on additions and multiplications (Imbo et al., 2007; Núñez-Peña et al., 2011; Smith-Chant & LeFevre, 2003) already showed a stronger problem-size effect in individuals with low arithmetic skill than in individuals with high arithmetic skill. Our behavioral results extend the previous findings to subtraction by demonstrating that the problem-size effect also depends on arithmetic ability in these less-practiced arithmetical problems whose poorer performance is often attributed to a greater use of procedures (Barrouillet et al., 2008; LeFevre et al., 2006; Seyler et al., 2003). Thus, we found that the increase in response time and the drop in hit rate from small to large subtractions were higher in less-skilled individuals than in their highly skilled peers.

According to Campbell and Xue (2001), there are three possible sources for the problem-size effect and discerning between them can help us to broaden our understanding of the effects of ability on the use of strategies. The first is that highly skilled individuals may use retrieval in large size problems more often than their less-skilled peers (LeFevre et al., 1996). The second is that highly skilled individuals use the retrieval strategy more efficiently in these problems. The third is similar to the second, only that in this case the proposal is that highly skilled individuals use procedures more efficiently than their less-skilled counterparts.

Our results rule out the first possibility, that is, that there are differences in the proportion of retrieval and procedure use depending on ability and problem size, since no difference was found in the frequency of reported retrieval between HS and LS participants. Interestingly, this corroborates the results reported by Thevenot et al. (2010). Moreover, our data are similar to those obtained in previous research (Campbell &

Xue, 2001; LeFevre et al., 2006). Individuals used retrieval in 70% of the small subtractions (82% in LeFevre et al. and 73% in Campbell and Xue), and in 49% of the large subtractions (42% both in LeFevre et al. and in Campbell and Xue).

As for the second possible source of individual differences in the problem-size effect, our results suggest that these differences are not due to more efficient retrieval in high-ability arithmetic problem solvers than in their less-skilled peers since groups did not differ in response time when they reported having used retrieval. Finally, we studied the third possible source of individual differences, namely differences in the efficiency of procedures depending on arithmetic skill. Our data supported this hypothesis: high- and low-ability individuals did not differ in the amount of use of procedures reported, but the higher skilled participants were faster and less error-prone than the less skilled ones when they reported having used this strategy.

The results of the present experiment also shed light on the findings reported by Thevenot et al. (2010). They found a discrepancy between the information obtained from verbal reports and that obtained by the operand-recognition paradigm concerning the strategy selected to solve subtractions. The results of the operand-recognition paradigm led Thevenot and colleagues to conclude that the selection strategy depends on arithmetic ability and problem size. Their data suggested that high- and low-ability arithmetic problem solvers do not differ in strategy selection for small subtractions (retrieval), but that highly skilled individuals rely mainly on retrieval when solving their medium-sized subtractions (i.e. our large ones) while their less-skilled peers mainly used procedures. In their study, Thevenot et al. also contrasted the results obtained with the operand-recognition paradigm and those obtained with the classical method of verbal report collection, and found that strategy reports did not differ between high- and low-ability groups with any subtraction size. Retrieval was mainly reported for small subtractions and in the case of medium subtractions, low-ability participants reported using procedures in 50% of the trials and their high-ability peers reported using procedures in 55%, suggesting that the two groups used retrieval and procedures equally to solve them. This result seemed to contradict the one obtained with the operand-recognition paradigm and led the authors to dismiss the report data as unreliable.

However, this discrepancy could be explained in terms of differences in the efficiency of procedures used by high- and low-ability individuals. According to our results, high-ability individuals' recognition times might not have differed after the calculation task and the comparison task with medium-sized numbers

(i.e. our large ones) because they used more efficient procedural strategies. The use of efficient calculation strategies might avoid the impairment in memory of the traces of the operands involved in the problem, and therefore the time taken to recognize the operands would not differ after a calculation task and after a comparison task. Additionally, participants with different skills might have used different procedures. As an example, an efficient procedure known as *addition reference* can be used to solve a subtraction, by directly recovering the addition from long-term memory (i.e., $12 - 4 = 8$ because $4 + 8 = 12$). In fact, knowledge of the inverse relation between addition and subtraction means that skill in addition improves performance in subtraction (Campbell, 2008; LeFevre et al., 2006; Seyler et al., 2003). Overall, differences in performance and recognition reaction time between groups in medium subtractions may be due to differences in the selection of procedural strategies. This hypothesis is also supported by the recent findings by Fayol and Thevenot (2012; see also Barrouillet and Thevenot, 2013) indicating that expert arithmetic solvers replace inefficient procedures such as counting by principled procedural knowledge which allows them to reach the solution more quickly and accurately (although see Campbell & Beech, 2014, for recent evidence that questions the use of automated procedures for simple addition in a sample of diversely-skilled participants). In any case, this alternative explanation complicates the application of the operand recognition paradigm as a precise measure of strategy selection in arithmetic, because it may not distinguish between efficient (fast) retrieval and equally efficient (fast) procedures. This criticism of the interpretation of the operand recognition paradigm can be added to those put forward by Metcalfe and Campbell (2010, 2011), who indicated that findings for performance in the operand recognition paradigm may also be modulated by task-switching and difficulty-related carryover effects.

Verbal reports like the ones used in this study have also their limitations, though. Fayol and Thevenot (2012)'s research was inspired by Baroody (1983, 1984) who claimed that in time the sophisticated procedures used by expert solvers would become too fast and automatic to reach consciousness. If this was the case, it might be that some of the trials in which our participants reported having used retrieval were really solved through automatic, unconscious procedures. Although we acknowledge this possibility, this does not impact our main result, which is that HS and LS individuals differed when they reported having used procedures.

A second aspect that requires further consideration is the fact that problems in our study were not presented as a whole but sequentially. This type of presentation might have supposed an extra load in the participants'

working memory, and affected the kind of strategies selected. We think that this possibility is highly unlikely, given that the percentages of strategies' use were quite similar to those reported in previous studies with standard presentation (Campbell & Xue, 2001; LeFevre et al., 2006). However, we cannot completely dismiss this possibility without replicating our data in a more ecological setting.

In conclusion, the current study broadens our understanding of the sources of individual differences in the problem-size effect by investigating a less-practiced operation such as subtraction. Whereas previous research associated arithmetical skill with an increase in the use of retrieval, recent studies have shown that highly skilled problem solvers also use procedures more efficiently (Barrouillet & Thevenot, 2013; Campbell et al., 2013; Fayol & Thevenot, 2012). Our finding extends this recent evidence to subtraction by showing that highly skilled individuals apply faster and less error-prone procedures than their less highly skilled counterparts. Furthermore, our results shed light on the discrepancies found previously between verbal reports and behavioral data, and confirm the need to support latency and error rate analyses with other measures.

References

- Ashcraft M. H., & Christy K. S. (1995). The frequency of arithmetic facts in elementary texts: Addition and multiplication in grades 1–6. *Journal for Research in Mathematics Education*, 26, 396–421. <http://dx.doi.org/10.2307/749430>
- Ashcraft M. H., & Guillaume M. M. (2009). Mathematical cognition and the problem size effect. In B. Ross (Ed.), *The psychology of learning and motivation* (pp. 121–151). Burlington, VT: Academic Press.
- Baroody A. J. (1983). The development of procedural knowledge: An alternative explanation for chronometric trends of mental arithmetic. *Developmental Review*, 3, 225–230. [http://dx.doi.org/10.1016/0273-2297\(83\)90031-X](http://dx.doi.org/10.1016/0273-2297(83)90031-X)
- Baroody A. J. (1984). A reexamination of mental arithmetic models and data: A reply to Ashcraft. *Developmental Review*, 4, 148–156. [http://dx.doi.org/10.1016/0273-2297\(84\)90004-2](http://dx.doi.org/10.1016/0273-2297(84)90004-2)
- Barrouillet P., & Thevenot C. (2013). On the problem-size effect in small additions: Can we really discard any counting-based account? *Cognition*, 128, 35–44. <http://dx.doi.org/10.1016/j.cognition.2013.02.018>
- Barrouillet P., Mignon M., & Thevenot C. (2008). Strategies in subtraction problem solving in children. *Journal of Experimental Child Psychology*, 99, 233–251. <http://dx.doi.org/10.1016/j.jecp.2007.12.001>
- Campbell J. I. D. (2008). Subtraction by addition. *Memory and Cognition*, 36, 1094–1102. <http://dx.doi.org/10.3758/MC.36.6.1094>
- Campbell J. I. D., & Alberts N. M. (2009). Operation-specific effects of numerical surface form on arithmetic strategy. *Journal of Experimental Psychology: Learning, Memory and*

- Cognition*, 35, 999–1011. <http://dx.doi.org/10.1037/a0015829>
- Campbell J. I. D., & Beech L. C.** (2014). No generalization of practice for non-zero simple addition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. <http://dx.doi.org/10.1037/xlm0000003>
- Campbell J. I. D., & Xue Q.** (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General*, 130, 299–315. <http://dx.doi.org/10.1037/0096-3445.130.2.299>
- Campbell J. I. D., Chen Y., & Maslany A. J.** (2013). Retrieval-induced forgetting of arithmetic facts across cultures. *Journal of Cognitive Psychology*, 25, 759–773.
- Cooney J. B., & Ladd S. F.** (1992). The influence of verbal protocol methods on children's mental computation. *Learning and Individual Differences*, 4, 237–257. [http://dx.doi.org/10.1016/1041-6080\(92\)90004-X](http://dx.doi.org/10.1016/1041-6080(92)90004-X)
- Duverne S., & Lemaire P.** (2005). Arithmetic split effects reflect strategy selection: An adult age comparative study in addition comparison and verification tasks. *Canadian Journal of Experimental Psychology*, 59, 262–278. <http://dx.doi.org/10.1037/h0087479>
- El Yagoubi R., Lemaire P., & Besson M.** (2003). Different brain mechanisms mediate two strategies in arithmetic: Evidence from event-related brain potentials. *Neuropsychologia*, 41, 855–862. [http://dx.doi.org/10.1016/S0028-3932\(02\)00180-X](http://dx.doi.org/10.1016/S0028-3932(02)00180-X)
- El Yagoubi R., Lemaire P., & Besson M.** (2005). Effects of aging on arithmetic problem-solving: An event-related brain potential study. *Journal of Cognitive Neuroscience*, 17, 37–50. <http://dx.doi.org/10.1162/08998929052880084>
- Fayol M., & Thevenot C.** (2012). The use of procedural knowledge in simple addition and subtraction problems. *Cognition*, 123, 392–403. <http://dx.doi.org/10.1016/j.cognition.2012.02.008>
- French J. W., Ekstrom R. B., & Price L. A.** (1963). *Manual for kit of reference tests for cognitive factors*. Princeton, NJ: Educational Testing Service.
- Grabner R. H., Ansari D., Koschutnig K., Reishofer G., Ebner F., & Neuper C.** (2009). To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving. *Neuropsychologia*, 47, 604–608. <http://dx.doi.org/10.1016/j.neuropsychologia.2008.10.013>
- Grabner R. H., & De Smedt B.** (2011). Neurophysiological evidence for the validity of verbal strategy reports in mental arithmetic. *Biological Psychology*, 87, 128–136. <http://dx.doi.org/10.1016/j.biopsycho.2011.02.019>
- Groen G. J., & Parkman J. M.** (1972). A chronometric analysis of simple addition. *Psychological Review*, 79, 329–343. <http://dx.doi.org/10.1037/h0032950>
- Imbo I., & Vandierendonck A.** (2008a). Effects of problem size, operation, and working memory span on simple-arithmetic strategies: Differences between children and adults. *Psychological Research*, 72, 331–346. <http://dx.doi.org/10.1007/s00426-007-0112-8>
- Imbo I., & Vandierendonck A.** (2008b). Practice effects on strategy selection and strategy efficiency in simple arithmetic. *Psychological Research*, 72, 528–541. <http://dx.doi.org/10.1007/s00426-007-0128-0>
- Imbo I., Vandierendonck A., & Rosseel Y.** (2007). The influence of problem features and individual differences on strategic performance in simple arithmetic. *Memory & Cognition*, 35, 454–463. <http://dx.doi.org/10.3758/BF03193285>
- Kirk R. E.** (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, 56, 746–759. <http://dx.doi.org/10.1177/0013164496056005002>
- Kirk E. P., & Ashcraft M. H.** (2001). Telling stories: The perils and promise of using verbal reports to study math strategies. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 27, 157–175. <http://dx.doi.org/10.1037/0278-7393.27.1.157>
- LeFevre J. A., DeStefano D., Penner-Wilger M., & Daley K. E.** (2006). Selection of procedures in mental subtraction. *Canadian Journal of Experimental Psychology*, 60, 209–220. <http://dx.doi.org/10.1037/cjep2006020>
- LeFevre J. A., & Kulak A. G.** (1994). Individual differences in the obligatory activation of addition facts. *Memory and Cognition*, 22, 188–200. <http://dx.doi.org/10.3758/BF03208890>
- LeFevre J. A., Sadesky G. S., & Bisanz J.** (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22, 216–230. <http://dx.doi.org/10.1037/0278-7393.22.1.216>
- LeFevre J. A., Shanahan T., & DeStefano D.** (2004). The tie effect in simple arithmetic: An access-based account. *Memory and Cognition*, 32, 1019–1031. <http://dx.doi.org/10.3758/BF03196878>
- Metcalf A. W. S., & Campbell J. I. D.** (2010). Switch costs and the operand-recognition paradigm. *Psychological Research*, 74, 491–498. <http://dx.doi.org/10.1007/s00426-009-0272-9>
- Metcalf A. W. S., & Campbell J. I. D.** (2011). Adults' strategies for simple addition and multiplication: Verbal self-reports and the operand recognition paradigm. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 37, 661–672. <http://dx.doi.org/10.1037/a0022218>
- Núñez-Peña M. I., Gracia-Bafalluy M., & Tubau E.** (2011). Individual differences in arithmetic skill reflected in event-related brain potentials. *International Journal of Psychophysiology*, 80, 143–149. <http://dx.doi.org/10.1016/j.ijpsycho.2011.02.017>
- Russo J. E., Johnson E. J., & Stephens D. L.** (1989). The validity of verbal protocols. *Memory and Cognition*, 17, 759–769. <http://dx.doi.org/10.3758/BF03202637>
- Seyler D. J., Kirk E. P., & Ashcraft M. H.** (2003). Elementary subtraction. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 29, 1339–1352. <http://dx.doi.org/10.1037/0278-7393.29.6.1339>
- Siegler R. S., Adolph K. E., & Lemaire P.** (1996). Strategy choices across the life span. In L. R. Reder (Ed.), *Implicit memory and metacognition* (pp. 796–121). Mahwah, NJ: Erlbaum.
- Smith-Chant B. L., & LeFevre J. A.** (2003). Doing as they told and telling it like it is: Self-reports in mental arithmetic. *Memory and Cognition*, 31, 516–528. <http://dx.doi.org/10.3758/BF03196093>

- Thevenot C., Barrouillet P., & Fayol M.** (2001). Algorithmic solution of arithmetic problems and operands-answer associations in long-term memory. *Quarterly Journal of Experimental Psychology*, 52A, 599–611. <http://dx.doi.org/10.1080/02724980042000291>
- Thevenot C., Castel C., Fanget M., & Fayol M.** (2010). Mental subtraction in high- and lower skilled arithmetic problem solvers: Verbal reports versus operand-recognition paradigms. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 36, 1242–1255. <http://dx.doi.org/10.1037/a0020447>
- Thevenot C., Fanget M., & Fayol M.** (2007). Retrieval or non-retrieval strategies in mental addition? An operand-recognition paradigm. *Memory and Cognition*, 35, 1344–1352.
- 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. 331–346). New York, NY: Psychology Press.