

Practice effects on strategy selection and strategy efficiency
in simple mental arithmetic

Ineke Imbo & André Vandierendonck

Ghent University

Running head: Practice effects in simple arithmetic

Correspondence Address: Ineke Imbo

Department of Experimental Psychology

Ghent University

Henri Dunantlaan 2

B – 9000 Ghent

Belgium

Tel: +32 (0)9 2646409

Fax: + 32 (0)9 2646496

E-mail: Ineke.Imbo@UGent.be

Abstract

Two experiments were conducted to investigate the effects of practice on strategy selection and strategy efficiency in mental arithmetic. Participants had to solve simple addition or multiplication problems, after having received 0, 3, or 6 practice sessions (Experiment 1), and before and after having received 3 practice sessions (Experiment 2). Strategy selection was measured by means of trial-by-trial strategy reports, whereas strategy efficiency was measured by means of response latencies. Results showed significant practice effects on retrieval frequency, procedural frequency, retrieval efficiency, and procedural efficiency. However, practice effects on strategy efficiency appeared to be both strategy-specific (i.e., only for procedural strategies) and operation-specific (i.e., only for multiplication problems). Implications of the present results for mathematic cognition and its modeling are discussed.

Practice effects on strategy selection and strategy efficiency
in simple mental arithmetic

Daily, we use several numeric competencies, such as subitizing small quantities, estimating large quantities, calculating new quantities, et cetera. Some of these basic competencies, such as subitizing and estimating, seem to be innate to human infants (e.g., Butterworth, 1999; Dehaene, 1997; Spelke & Dehaene, 1999). The mastery of more advanced numerical skills such as calculation, in contrast, must be acquired through education, learning, and practice¹. Between the ages of 2 and 4, children learn to count verbally (e.g., Gelman & Gallistel, 1978; Wynn, 1990). Once they master counting, children generalize the counting procedure to larger numbers, apparently without upper bound and without explicit training. During their elementary school years, children learn a set of basic arithmetic facts and calculation procedures. Slow, deliberate and effortful procedures (such as step-by-step counting) are replaced by fast, efficient, and less effortful calculation processes (such as memory retrieval; e.g., Ashcraft, 1982; Fuson, 1982, 1988; Siegler, 1988; Steel & Funnell, 2001). When reaching adult age, until recently, people were supposed always to use memory retrieval to solve simple-arithmetic problems such as $3 + 5$ and 6×7 (e.g., Ashcraft, 1992; McCloskey, 1992). However, more recent studies showed that this is not absolutely true: even skilled adults are not always able to retrieve simple-arithmetic facts from their memory (e.g., LeFevre, Bisanz, et al., 1996a; LeFevre, Sadesky, & Bisanz, 1996b). Indeed, many adults use nonretrieval (procedural) strategies such as counting (e.g., $6 + 3 = 6 + 1 + 1 + 1$) and transformation (e.g., $6 + 7 = 6 + 4 + 3$) to solve simple-arithmetic problems. Other quite surprising results have been found concerning the arithmetic abilities of normal educated adults. Geary and colleagues (1996, 1997) showed declines in mental arithmetic

performance across successive North-American generations. Comparable results were obtained by Mulhern and Wylie (2004), who showed that performance levels of British psychology students on core mathematical skills (such as calculation) dropped devastatingly between 1992 and 2002.

The importance of mental arithmetic in daily life, and the decline of mathematic skill the last few years notwithstanding, few studies so far investigated the effects of practice on strategy selection and strategy efficiency. Practice effects in simple arithmetic have been studied in children (see Kroesbergen & Van Luit, 2003, for a review), and in brain-damaged adults (e.g., Whetstone, 1998), but nearly not in healthy adults (but see Fendrich, Healy, & Bourne, 1993; Pauli, Bourne, & Birbaumer, 1998; Rickard, Healy, & Bourne, 1994). Four questions may be raised: (1) Does practice increase the use of direct memory retrieval (i.e., a change in strategy selection)?, (2) Does practice increase the speed with which retrieval and procedural strategies are executed (i.e., an increase in strategy efficiency)?, (3) Does practice reduce the performance differences between small and large problems (i.e., the problem-size effect)?, and (4) Do practice effects transfer to other operations, other sizes, or other arithmetic problems?

Concerning the first question (i.e., whether practice influences strategy selection), the most prevailing assumption is that practice will inevitably lead to an augmented usage of retrieval. This assumption is based on the distribution of associations model (Siegler & Shrager, 1984) and the instance theory of automatization (Logan, 1988). The distribution of associations model (Siegler & Shrager, 1984) states that the encoding of a problem results in the activation of a set of response candidates. The activation of each candidate depends on the acquired problem-answer strength. It is further assumed that there is a direct relation between the activation level and the probability of retrieval. Answers with a high associative strength will be retrieved, but if

the problem-answer associative strength does not exceed a predefined confidence criterion, a procedural strategy will be used in order to solve the problem. As continued practice strengthens problem-answer associations, retrieval will be used more frequently, resulting in a concomitant decrease in procedural strategy use. In a later version of this model, the adaptive strategy choice model² (ASCM, Siegler & Shipley, 1995), selection of an arithmetic strategy depends on its relative efficiency (i.e., speed and accuracy). As a result of practice, problem-strategy associations increase, and this increase is as large for both retrieval and procedural strategies. This model thus not only predicts an increase in the use of retrieval strategies, but also an increase in the use of more efficient procedural strategies (e.g., transformation) relative to less efficient ones (e.g., counting). However, in the end, extensive practice should result in exclusive retrieval use.

According to Logan's (1988) instance theory of automatization, each encounter with a stimulus initiates a race between procedural and retrieval strategies. In the beginning, the race is predominantly won by procedures. However, as each problem encounter is encoded and stored in long-term memory, practice enhances retrieval speed but not procedural speed. Consequently, as practice progresses, the retrieval strategy will win the race. Otherwise stated, practice enhances the amount of automatization, which reflects a transition from performance based on procedural strategies to performance based on memory retrieval.

However, since the use of procedural strategies persists even in skilled adults, LeFevre et al. (1996a) maintain that practice will *not always* lead to increased usage of retrieval. According to these authors, practice can also lead to the automatic activation of procedural strategies. In this view, associations between a specific problem and a procedure are created and strengthened by the successful use of such procedural strategies. Therefore, when people encounter that problem,

they will automatically activate a procedure to solve that problem, without (or before) trying to retrieve it from long-term memory. Practice will then not solely lead to the replacement of procedures by fact retrieval, but also to the replacement of less efficient procedures by more efficient procedures. This view (see also Baroody, 1983, 1984, 1985) implies that the availability of efficient procedural skills would avoid the necessity of memorizing all the basic number combinations. Consequently, many simple-arithmetic problems might continue to be solved by using procedural strategies.

Up until now, evidence concerning practice effects on strategy selection has only been shown for the alphabet arithmetic task (e.g., Brigman & Cherry, 2002; Compton & Logan, 1991; Hoyer, Cerella, & Onyper, 2003; Logan, 1988; Logan & Klapp, 1991, Rickard, 2004; Zbrodoff, 1999), for pseudo-arithmetic tasks (e.g., Onyper, Hoyer, & Cerella, 2006; Rickard, 1997; Touron, Hoyer, & Cerella, 2004), and for arithmetic word problems (e.g., Lewis, 1989). The current study aims to contribute to this literature by investigating practice effects on strategy selection in simple-arithmetic tasks.

The second question concerns practice effects on strategy efficiency. According to the distribution of associations model (Siegler & Shrager, 1984), successive correct practice trials strengthen the link between a problem and its answer. Since the time to retrieve and produce an answer is proportional to the activation level of the corresponding answer node, this theory predicts that practice will increase retrieval efficiency. The later version of this model, the ASCM (Siegler & Shipley, 1995), predicts an increase in the efficiency with which *each* strategy is executed, and thus predicts higher efficiencies in both retrieval and procedural strategies. Similar predictions can be made based on Rickard's (1997) theory of skill acquisition, the component power law (CMPL). This theory explains increases in procedural efficiency as a function of

increases in retrieval efficiency. A computation is first reduced to its simpler parts, each of which is resolved by direct memory retrieval (see also Anderson, 1993). Hence, both retrieval and procedural efficiency increase when memory traces are strengthened.

The instance theory of automatization (Logan, 1988), in contrast, assumes that the finishing times for procedures stay the same while the finishing times for retrieval decrease. This theory thus predicts practice effects on retrieval efficiency but not on procedural efficiency. Finally, Baroody (1983, 1984, 1985) predicts rather the opposite. According to his procedure-based theory, skill acquisition is based on the replacement of slow procedural strategies by faster, more automatic procedural strategies rather than the replacement of procedures by direct memory retrieval. He thus predicts stronger practice effects on procedural efficiency than on retrieval efficiency.

Effects of practice on strategy efficiency have been observed in standard arithmetic problems (e.g., Campbell, 1987, 1999; Fendrich et al., 1993; Pauli et al., 1998; Rickard, 2005; Rickard & Bourne, 1996; Rickard et al., 1994, Royer, Tronsky, Chan, Jackson, & Marchant, 1999), in pseudo-arithmetic tasks (e.g., Onyper et al., 2006; Rickard, 1997; Tournon et al., 2004), and in alphabet arithmetic tasks (e.g., Brigman & Cherry, 2002; Compton & Logan, 1991; Hoyer et al., 2003; Klapp, Boches, Trabert, & Logan, 1991; Logan, 1988; Logan & Klapp, 1991; Rickard, 2004). However, since most of these studies did not include trial-by-trial strategy reports, it is not clear whether the faster response times were due to changes in strategy efficiency (i.e., faster strategy execution) or to changes in strategy selection (i.e., more frequent use of faster strategies). Compton and Logan (1991) and Rickard (1997, 2004) included strategy reports on *subsets* of trials (e.g., on one sixth of the trials) whereas Logan and Klapp (1991) asked participants – at the end of the experiment – to estimate the percentage of trials on which they

had used retrieval vs. counting strategies. Strategy reports on *all* trials have been used in alphabet arithmetic tasks (e.g., Hoyer et al., 2003) and in pseudo arithmetic tasks (e.g., Onyper et al., 2006; Tournon et al., 2004), but not yet in ‘pure’ simple-arithmetic tasks. In the present study, trial-by-trial strategy reports are used to investigate practice effects on retrieval efficiency and procedural efficiency separately.

Thirdly, the present study aims to test why the problem-size effect is modified by practice. The problem-size effect refers to the observation that large problems such as 8×9 take longer to solve than small problems such as 2×3 . The problem-size effect decreases as a result of practice, both in ‘pure’ arithmetic tasks (e.g., Fendrich et al., 1993; LeFevre & Liu, 1997; Pauli et al., 1998; Rickard & Bourne, 1996) and in the alphabet arithmetic task (e.g., Brigman & Cherry, 2002; Logan, 1988; Zbrodoff, 1995). However, practice never eliminated the problem-size effect, even at asymptotic response times. The problem-size effect may decrease in three ways: more frequent retrieval use for large problems, more efficient retrieval use for large problems, and more efficient procedural use for large numbers (Campbell & Xue, 2001). The present study tests whether practice influences all three of them.

The final question raised, concerns the transfer of practice to other operations, other sizes, or other arithmetic problems. According to associative network theories (e.g., Campbell, 1987; Campbell & Graham, 1985), the instance theory of automatization (Logan, 1988), and the identical elements model (Rickard, 2005; Rickard et al., 1994), practice effects are item-based rather than process-based. This implies that practice involves learning specific responses to specific stimuli. Consequently, transfer to novel stimuli and situations should be inexistent.

According to the ASCM (Siegler & Shipley, 1995) and the CMPL theory (Rickard, 1997), in contrast, practice enhances both retrieval and procedural efficiencies. If it is further assumed

that procedures can be applied to several stimuli, these models predict that practice effects on simple-arithmetic problems will transfer to complex-arithmetic problems. This reasoning is also adopted in the procedure-based view of Baroody (1983, 1984, 1985), which implies that procedural strategy use is cognitively more economical than retrieval use because it can be used on multiple problems.

Previous studies on mental arithmetic reported transfer for highly related problems (e.g., commuted problems, but not for other problems or other operations (e.g., Campbell, 1987a; Fendrich, et al., 1993; Pauli et al., 1994; Rickard & Bourne, 1996; Rickard et al., 1994). Pauli et al. (1998) even did observe no overall transfer from practiced to new multiplication problems. Practice was thus item-specific and did not facilitate arithmetic performance on problems that were not practiced. More recently, Delazer et al. (2005) showed that transfer from old to new complex addition problems only occurred when procedural strategies had been practiced but not when direct memory retrieval had been practiced. We wondered whether the same strategy-dependent effect of transfer would be true when simple arithmetic is practiced.

The present study consists of two experiments, which were conducted in order to formulate an answer to the four questions outlined above.

Experiment 1

In this experiment, participants had to solve simple-arithmetic addition or multiplication problems. On the assumption that small problems are usually solved more efficiently than are large problems (cf. the problem-size effect), only large problems were practiced. Participants

were given 0, 3, or 6 practice sessions on the subset of largest problems. After these practice sessions, a test session was administered, in which both small and large problems had to be solved. The latency data and strategy reports collected in the test session were used to investigate practice effects on strategy efficiency and strategy selection, respectively. After the test session, a test of complex arithmetic (the French kit) was administered as well.

Method

Participants. Sixty first-year psychology students (9 men and 51 women) at Ghent University participated for course requirements and credits. Their mean age was 19.0 years. They were randomly assigned to the cells of a 2 (Operation) x 3 (Practice) design.

Procedure. Each participant was tested individually in a quiet room for approximately 30, 45, or 60 minutes (dependent on the practice condition). Two tasks were given to each participant. The first one was the simple-arithmetic task, which consisted of simple additions (for one group of 30 participants) or simple multiplications (for another group of 30 participants). Within each group, 10 participants did not practice, 10 participants completed 3 practice sessions, and 10 participants completed 6 practice sessions. The test session was administered after the practice sessions. The second task was the French kit (French, Ekstrom, & Price, 1963), which consists of complex-arithmetic problems that have to be solved as quickly and accurately as possible. In the following, both tasks are described more extensively.

Stimuli of the simple-arithmetic task consisted of simple addition and simple multiplication problems. Both addition and multiplication problems were composed of pairs of

numbers between 2 and 9, with tie problems (e.g., $3 + 3$) excluded. Problems involving 0 or 1 as an operand or answer were also excluded. This resulted in 56 addition problems (ranging from $2 + 3$ to $8 + 9$) and 56 multiplication problems (ranging from 2×3 to 8×9). Although all problems were presented in the test session, only the most difficult problems were presented in the practice sessions. The practice problems consisted of the 12 largest addition problems and the 12 largest multiplication problems. For addition, this selection included all problems with a sum ranging from 14 to 17. For multiplication, this selection included all problems with a product ranging from 45 to 72. Definition of small and large problems was also based on this selection: small problems were defined as the not-selected problems (i.e., the 44 smallest ones), whereas large problems were defined as the selected problems (i.e., the 12 largest ones). As noted before, there were three practice conditions: 0, 3, or 6 practice sessions. Within each practice session, all practice problems (i.e., the large ones) were presented twice, and in the test session all problems (small and large ones) were presented twice. All problems were presented in Arabic format and in a randomized order within one session.

A trial started with a fixation point, which appeared for 500 msec. Then the arithmetic problem appeared horizontally in the center of the screen, with the operation sign at the fixation point. The problem remained on screen until the participant responded. In order to avoid biasing conditions, no time deadline was set, since it has been shown that a fast deadline increases reported use of retrieval, especially for large problems (Campbell & Austin, 2002). A sound-activated relay was activated when participants spoke their answer aloud in a microphone, which was connected to a software clock (accurate to 1 msec). The use of a voice-key minimized general speeding effects in motor responses during practice. In previous research (e.g., Rickard et al., 1994) participants often had to type in the answer on the numeric key pad, so that

improvements in motor aspects during practice might have influenced overall performance. Pauli et al. (1998) indeed showed decreases in both mental calculation time and motor response time across practice sessions. All invalid trials (e.g., failures of the voice-activated relay) were discarded, and (in the test session only) they returned at the end of the session.

On each trial, accuracy was registered online by the experimenter and feedback was presented to the participants, a green 'Correct' when their answer was correct, and a red 'Incorrect' when it was not. Participants were also told to report the strategy they used for each single problem. The reported strategy was recorded online by the experimenter by pressing a predefined number key on the keyboard. Participants could choose one of the four strategies described below (see e.g., Campbell & Gunter, 2002; Campbell & Xue, 2001; Kirk & Ashcraft, 2001; LeFevre et al., 1996b; Seyler, Kirk, & Ashcraft, 2003): (1) *Remember: You solve the problem by just remembering or knowing the answer directly from memory;* (2) *Counting: You solve the problem by counting a certain number of times to get the answer;* (3) *Transformation: You solve the problem by referring to related operations or by deriving the answer from some known facts;* and (4) *Other: You solve the problem by a strategy unlisted here, or you do not know what strategy you used to solve the problem.* These four strategies were extensively explained by the experimenter, with examples of both addition or multiplication problems solved by each strategy as appropriate. It was emphasized that the presented strategies were not meant to encourage use of a particular strategy.

After the simple-arithmetic task, participants completed two arithmetic subtests of the French kit (French et al., 1963), one page of complex addition problems (e.g., $39 + 90 + 82$) and one page of complex subtraction and multiplication problems (e.g., 48×7). Each page contained six rows of ten vertically oriented problems. Participants were given two minutes per page to

solve the problems as quickly and accurately as possible. Scores were defined as the total number of correctly solved problems per test.

Results

Across operations, 6.8% of the trials was spoiled due to failures of the sound-activated relay. Since all the invalid trials met in the test session returned at the end of this session, most of them were recovered from data loss, which reduced the trials due to failures of the sound-activated relay to 0.5%. Further, all incorrect trials (3.0%) and all trials on which participants selected the 'Other' category (0.1%) were deleted. All response times (RTs) more than 4 standard deviations from each participant's mean (per operation) were discarded as outliers (0.7%). Finally, one participant (in the multiplication experiment with six practice sessions) was discarded due to voice key problems. Every analysis of variance (ANOVA) that has been executed was based on the multivariate linear model. All reported results are considered to be significant if $p < .05$, unless mentioned otherwise. Although no pre-practice test had been administered, we may assume that all observed effects were due to the manipulated variables, given that the RTs in the first practice session did not differ across groups (0x, 3x, 6x; $F < 1$)³.

Strategy selection. Practice effects on strategy selection were tested by means of three separate ANOVAs – one for each strategy type (retrieval, transformation, and counting; see Table 1). These ANOVAs were run on percentages of strategy use in the test session, with problem size (small vs. large) as within-subjects variable and practice level (0x, 3x, 6x) and operation (addition vs. multiplication) as between-subjects variables. The ANOVA on percentages of retrieval use

showed no main effect of practice level, $F(2,53) = 2.11$ ($p = .13$), but planned comparisons showed that 6x practiced participants used the retrieval strategy more often than did 3x practiced participants, $F(1,53) = 4.04$, whereas there was no difference between the 0x and 3x practiced participants, $F < 1$. Conversely, the ANOVA on percentages of counting use showed that 6x practiced participants used the counting strategy *less* often than did 3x practiced participants, $F(1,53) = 5.49$. Similarly, there was no difference in counting use between the 0x and 3x practiced participants, $F < 1$. The main effect of practice level was insignificant in the ANOVA on percentages of transformation use ($F < 1$).

Though, in all three ANOVAs, session interacted with problem size, $F(2,53) = 7.94$ for retrieval, $F(2,53) = 2.44$ ($p = .09$) for transformation, and $F(2,53) = 9.35$ for counting. For small (unpracticed) problems, the amounts of retrieval, transformation, and counting use were equally high across the 0x, 3x, and 6x practiced participants (all F s < 1). For large (practiced) problems, in contrast, the amount of retrieval use significantly *increased* across practice levels, $F(1,53) = 6.25$ whereas the amount of counting use significantly *decreased*, $F(1,53) = 11.17$. The amount of transformation use on large problems also decreased, but this effect did not reach significance, $F(1,53) = 1.30$ ($p = .25$). Anyhow, even after practice, retrieval was still used more frequently on small problems (77%) than on large problems (60%), $F(1,53) = 39.19$. Importantly, the problem-size effect in terms of retrieval use (i.e., % retrieval use on small problems – % retrieval use on large problems) was significantly higher for the 0x practiced participants (31%) than for the 3x practiced participants (15%) and the 6x practiced participants (4%), $F(1,53) = 5.89$ and $F(1,53) = 15.58$, respectively.

Finally, we tested whether there were differences across the operations. Retrieval was used more frequently on multiplication (79%) than on addition (57%), $F(1,53) = 13.28$, whereas

transformation was used more frequently on addition (36%) than on multiplication (11%), $F(1,53) = 24.46$. Counting was used as frequently on addition (7%) and multiplication (10%), $F < 1$. Importantly, the effect of operation did not interact with any other variable in any of the three ANOVAs.

Strategy efficiency. In order to test practice effects on strategy efficiency, an ANOVA on RTs in the test session was performed with practice level (0x, 3x, 6x) and operation (addition vs. multiplication) as between-subjects variables, and problem size (small vs. large) and strategy (retrieval vs. procedural) as within-subjects variables (see Table 2)⁴. Obviously, the main effect of strategy reached significance, $F(1,53) = 124.17$; with procedural RTs (2033 msec) being higher than retrieval RTs (1021 msec). This result indicates that participants' verbal strategy reports are highly reliable. The main effect of practice level was significant as well, $F(2,53) = 6.07$. RTs were larger for the 0x practice level (1835 msec) than for both 3x and 6x practice levels (1300 msec and 1446 msec), $F(1,53) = 11.38$ and $F(1,53) = 5.85$, respectively. RTs did not differ between 3x and 6x practice levels, $F < 1$. Practice level interacted with strategy, $F(2,53) = 9.89$, and with operation, $F(2,53) = 4.43$. Practiced participants were more efficient in performing procedural strategies than unpracticed participants, $F(1,53) = 14.77$, whereas practiced and unpracticed participants were equally efficient in the retrieval strategy, $F < 1$. Furthermore, participants did not differ in efficiency on additions, $F < 1$, but practiced participants were more efficient than unpracticed participants on multiplications, $F(1,53) = 19.78$. The three-way interaction between practice level, strategy, and operation, $F(2,53) = 9.61$, confirmed that practiced participants were better than unpracticed participants in solving multiplication problems with procedural strategies, $F(1,53) = 27.79$, but not in solving multiplication problems via

retrieval, $F < 1$, nor in solving addition problems with retrieval or procedural strategies (each $F < 1$).

There was no interaction between problem size and practice level, $F(2,53) = 1.69$ ($p = .19$). Planned comparisons confirmed that the retrieval problem-size effect (i.e., retrieval RTs large problems – retrieval RTs small problems) was significant in all groups and did not differ across groups, neither for addition nor for multiplication (each $F < 1$). The procedural problem-size effect (i.e., procedural RTs large problems – procedural RTs small problems) for multiplication was significantly larger for the 0x practice group than for the 3x practice group, $F(1,53) = 5.55$, but equally large in the 3x and 6x practice groups, $F(1,53) = 1.72$ ($p = .20$). The procedural problem-size effect for addition did not differ across groups either (each $F < 1$). These results indicate that the effect of practice on the problem-size effect originates from more efficient procedural strategy use; an effect that seems to be reliable for multiplication only.

Transfer effects. As mentioned before, an additional test of complex-arithmetic performance (the French kit) was administered for each participant. Since this test was administered after the test session (and thus after the practice sessions for the practiced participants), scores on this test are suggestive for transfer effects of simple-arithmetic practice on complex-arithmetic performance. An operation (addition or multiplication) x practice level (0x, 3x, 6x) ANOVA was run for each subtest of the French Kit (i.e., the addition subtest and the subtraction-multiplication subtest). For the addition subtest, no significant effects appeared (each $F < 1$), indicating no transfer effects at all. For the subtraction-multiplication subtest, in contrast, both main effects were significant. The practiced participants (3x and 6x) performed significantly better than the unpracticed (0x) participants, $F(1,53) = 5.1$ (means of 17.0, 16.7, and 13.8, respectively), indicating a transfer effect from simple to complex problems. Participants who had

practiced multiplication problems performed significantly better than participants who had practiced addition problems, $F(1,55) = 5.5$ (means of 17.4 and 14.3, respectively), indicating that the transfer effect was operation specific.

Summary. Practice resulted in more frequent retrieval use, less frequent counting use, more efficient retrieval use, and more efficient procedural use. We may thus conclude that practice influenced both strategy selection and strategy efficiency. However, the effects on strategy efficiency were operation specific, as they were only apparent for multiplication and not for addition. Similarly, transfer effects from simple to complex problems were only significant for multiplication. The problem-size effect, finally, was reduced but did not completely disappear. Practice effects on the problem-size effect were associated with more frequent retrieval use and more efficient procedural use for large problems, but not with more efficient retrieval use for large problems. All these interesting observations notwithstanding, this experiment had one drawback: There was no pretest. Experiment 2 was meant to investigate the results obtained in Experiment 1 more thoroughly, by including both a pretest and a posttest.

Experiment 2

The present experiment differed from Experiment 1 in three aspects. First, Experiment 2 included both a pretest and a posttest whereas Experiment 1 only included a posttest. In the pretest, participants had to solve all (i.e., both small and large) problems once. In the following practice sessions, only the large problems were practiced. In the posttest, all problems (small and large ones) had to be solved again. Second, a speeded verification task was included. This task

was administered twice: once before the pretest and once after the posttest. This task was meant to test whether the practice effects on retrieval use observed in Experiment 1, were due to a real change in the retrieval network (i.e., the sensitivity) or to response biases. Indeed, trial-by-trial strategy reports have been criticized (e.g., Kirk & Ashcraft, 2001), since participants' strategy reports may easily be biased by the experimenter's suggestions. More specifically, participants might want to please the experimenter by reporting more frequent retrieval use without really using retrieval more frequently. As explained below, signal-detection theory can be used to disentangle real practice effects on strategy selection from effects caused by response biases. Third, since Experiment 1 showed that 3 practice sessions were enough to obtain sensitive differences with the control condition, the number of practice sessions was restricted to 3. All participants thus took part in a simple-arithmetic task consisting of a pretest, three practice sessions, and a posttest.

Method

Participants. Forty students (9 men and 31 women) at Ghent University participated in this Experiment. Half of them participated for course requirements and credits; the other half received €10 for participation. Their mean age was 20.0 years. None of them had participated in Experiment 1.

Procedure. Twenty participants completed simple addition problems and twenty participants completed simple multiplication problems. All participants also had to solve a complex-arithmetic test (the French kit) and to participate in a speeded verification task. The

basic procedure of this second experiment was identical to the one used in Experiment 1, except the three changes described above. Hence, only the aspects of the procedure that were different for this experiment (e.g., the inclusion of a speeded verification task) are described underneath.

In the speeded verification task, participants had to verify simple additions or simple multiplications, depending on the operation they had to solve in the simple-arithmetic production task. Stimuli of the speeded verification task were presented in standard form (i.e., $a + b = c$ or $a \times b = c$) in which a and b were one-digit numbers from 2 to 9. Half of the problems were presented with a correct solution whereas the other half were presented with an incorrect solution. The incorrect addition solutions were one or two units larger or smaller than the correct sum (e.g., $7 + 2 = 11$). The incorrect multiplication solutions were 10% or 20% larger or smaller than the correct product (e.g., $3 \times 4 = 10$). To reduce interference effects, stimuli were excluded when (a) $c = a \times b$ for addition problems (e.g., $3 + 2 = 6$) or $c = a + b$ for multiplication problems (e.g., $2 \times 3 = 5$), (b) $c = a$ or $c = b$ (e.g., $2 + 2 = 2$), (c) $c = N \times a$ or $N \times b$ for multiplication problems (e.g., $4 \times 5 = 16$), and (d) c is even (uneven) while the correct solution is uneven (even) (e.g., $3 \times 5 = 14$).

The verification task consisted of 8 practice trials and 80 experimental trials. A trial started with a fixation point for 500 msec, after which the stimulus was presented until the participants responded or until the response deadline was met. The response deadlines were based on the retrieval RTs of the 0x practice group in Experiment 1 and were calculated with the following formula: [*mean retrieval RT* + 2 * *standard deviation of retrieval RT*]. This measure was calculated separately for addition (1274 msec) and multiplication (1552 msec). After each practice trial, feedback was provided for one second, consisting of the word(s) “Correct” (when the answer was correct), “Incorrect” (when the answer was incorrect), or “Respond faster!” (when the participant’s response was slower than the response deadline). When the participant

answered within the response deadline, his/her response time appeared on the screen as well. No feedback was provided in the experimental trials, although participants were strongly recommended to answer as fast and accurately as possible. All RTs higher than the response deadlines were discarded. The inter-trial interval was 500 msec. The speeded verification task was administered twice: one before practice and once after practice.

Results

In the simple-arithmetic test, 7.0% of the trials were spoiled due to failures of the sound-activated relay. Since all the invalid trials met in the test session returned at the end of this session, most of them were recovered from data loss, which reduced the trials due to failures of the sound-activated relay to 0.5%. Further, all incorrect trials (3.5%) and all trials on which participants selected the 'Other' category (0.4%) were deleted. All RTs more than 4 standard deviations from each participant's mean (per operation) were discarded as outliers (0.7%).

Strategy selection. In order to test practice effects on strategy selection, we ran three separate ANOVAs – one for each strategy type (retrieval, transformation, and counting; see Table 3). These ANOVAs were run on percentages of strategy use, with problem size (small vs. large) and session (pre vs. post) as within-subjects variables, and operation (addition vs. multiplication) as between-subjects variable. Percentages of retrieval use were higher in the posttest (77%) than in the pretest (71%), $F(1,38) = 8.68$, whereas percentages of transformation use were lower in the posttest (20%) than in the pretest (26%), $F(1,38) = 8.79$. The amount of counting use did not change between the pretest (3%) and the posttest (3%), $F < 1$.

In all three ANOVAs, session interacted with problem size, $F(1,38) = 21.88$ for retrieval, $F(1,38) = 13.51$ for transformation, and $F(1,38) = 4.78$ for counting. For small (unpracticed) problems, the amounts of retrieval, transformation, and counting use stayed equally high across the sessions (all $ps > .10$). For large (practiced) problems, in contrast, the amount of retrieval use significantly *increased*, $F(1,38) = 15.13$ whereas the amounts of transformation and counting use significantly *decreased*, $F(1,38) = 49.63$ and $F(1,38) = 230.78$, respectively. Importantly, the problem-size effect in terms of retrieval use (i.e., % retrieval use on small problems – % retrieval use on large problems) was higher before practice (85% vs. 57%) than after practice (84% vs. 71%).

Finally, we tested whether there were differences across the operations. As in Experiment 1, retrieval was used more frequently on multiplication (85%) than on addition (63%), $F(1,38) = 13.43$, whereas transformation was used more frequently on addition (34%) than on multiplication (13%), $F(1,38) = 13.99$. Counting was used as frequently on addition (3%) and multiplication (2%), $F < 1$. Importantly, the effect of operation did not interact with any other variable in any of the three ANOVAs.

Strategy efficiency. Practice effects on strategy efficiency were tested with an ANOVA on RTs with problem size (small vs. large), session (pre vs. post) and strategy (retrieval vs. procedural) as within-subjects variables, and operation (addition vs. multiplication) as between-subjects variable (see Table 4 and footnote 4). Obviously, the main effect of strategy reached significance, $F(1,38) = 77.55$; with procedural RTs (1466 msec) being larger than retrieval RTs (917 msec). Participants thus seemed able to report reliably the strategies they used. RTs were faster in the posttest (1167 msec) than in the pretest (1216 msec), $F(1,38) = 4.08$. This was true

for both retrieval and procedural strategies, as appears from the insignificant session x strategy interaction, $F < 1$.

The main effect of problem size did not reach significance, $F(1,38) = 1.13$, but the interactions problem size x strategy and problem size x strategy x operation did, $F(1,38) = 25.93$ and $F(1,38) = 4.75$, respectively. The retrieval problem-size effect (i.e. retrieval RTs large problems – retrieval RTs small problems) was significant for both addition and multiplication, $F(1,38) = 15.65$ and $F(1,38) = 38.48$, respectively. Moreover, it did not change across sessions, $F < 1$ for addition and $F(1,38) = 1.64$ for multiplication. The procedural problem-size effect (i.e. procedural RTs large problems – procedural RTs small problems), in contrast, was not significant for addition, $F < 1$, but inverted for multiplication, $F(1,38) = 6.98$. More specifically, procedures were executed faster on large (i.e., practiced) multiplication problems than on small (i.e., unpracticed) multiplication problems. Obviously, this was only true in the posttest, $F(1,38) = 8.73$ and not in the pretest, $F < 1$. In conclusion, practice enhanced the procedural efficiency for multiplication but not for addition.

Transfer effects. As in Experiment 1, an additional test of complex-arithmetic performance (the French kit) was administered after the posttest. Separate ANOVAs were run for each subtest of the French Kit (i.e., the addition subtest and the subtraction-multiplication subtest) with operation (addition or multiplication) as the only independent variable. Participants having practiced simple additions scored slightly better on the complex-addition test than participants having practiced simple multiplications (15.6 vs. 14.9, respectively), but this effect did not reach significance, $F < 1$. The same was true for the complex-multiplication test, on which participants having practiced simple multiplications scored (insignificantly) better than participants having practiced simple additions (17.8 vs. 16.7, respectively), $F < 1$.

Speeded verification. The data of the speeded-verification task were analyzed with the help of the signal-detection theory in order to verify the validity of the participants' strategy reports. All trials on which participants had verified a correct addition or multiplication problem as correct were coded as hits. All trials on which participants had verified an incorrect addition or multiplication problem as correct were coded as false alarms. Using the software program of Van der Goten and Vandierendonck (1997), the signal-detection theory was used to determine the sensitivity (d') and the response bias (c) (see Table 5). A 2 x 2 ANOVA was conducted on these d' and c values, with operation (addition vs. multiplication) as between-subjects variable and session (pre vs. post) as within-subjects variable. The ANOVA on the sensitivity (d') showed a significant main effect of operation, indicating that the sensitivity was higher for multiplication (3.23) than for addition (2.08), $F(1,38) = 24.91$. Moreover, the increase in sensitivity (pre vs. post) tended to be significant for multiplication, $F(1,38) = 3.67$ ($p = .06$) but not for addition, $F < 1$. The ANOVA on response bias (c) showed no significant effects (highest $F = 2.99$, $p = .10$). The practice effects on strategy selection could thus be attributed to real differences rather than to changes in response biases.

Summary. As in Experiment 1, practice influenced both strategy selection and strategy efficiency. Indeed, direct memory retrieval was used more frequent after practice than before, transformation was used less frequent after practice than before, and retrieval and procedural use were more efficient after practice than before. Also as in Experiment 1, practice effects on strategy efficiency were larger for multiplication than for addition. The current experiment also confirmed that the problem-size effect was reduced by more frequent retrieval use and more efficient procedural use, but not by more efficient retrieval use. Finally, the speeded-verification task showed that the effects on strategy selection could not be attributed to response bias effects.

General Discussion

The present study revealed some remarkable findings concerning the effects of practice on simple-arithmetic performance. First, practice enhanced the frequency of retrieval use and reduced the frequency of procedural strategy use, and thus influenced *strategy selection*. Second, as retrieval RTs were lower after practice than before, *retrieval efficiency* was improved by practice as well. Third, practice augmented *procedural efficiency*, since procedural RTs decreased as a result of practice. Fourth, participants who had been practicing simple-arithmetic problems were better in a complex-arithmetic test than were unpracticed participants (cf. Experiment 1), which indicates transfer from simple to complex-arithmetic problems. However, this was not confirmed in Experiment 2. In the following, tentative answers to the four questions formulated in the introduction (i.e., practice effects on strategy selection, on strategy efficiency, on the problem-size effect, and transfer effects) are proposed on the basis of the present findings. We also check which arithmetic models are best fit to explain the results observed in this study.

Practice effects on strategy selection

Practice effects on strategy selection were significant in both experiments and for both operations. The increase in retrieval use was accompanied with a decrease in the use of counting strategies in Experiment 1, and with a decrease in the use of transformation strategies in Experiment 2. This difference across experiments is probably due to inter-individual differences.

The participants in Experiment 1 used the counting strategy rather frequently (cf. Table 1), and practice caused them to replace these counting strategies by transformation and/or retrieval strategies. As the participants in Experiment 2 used the counting strategy very rarely (cf. Table 3), they mainly switched from transformation use to direct memory retrieval.

However, the increase in reported retrieval use as a result of practice might be questioned, as it might be biased by demand effects (e.g., Kirk & Ashcraft, 2001). In that case, participants would (falsely) report more frequent retrieval use without really using retrieval more frequently. A signal detection analysis was used to disentangle effects caused by response bias and real practice effects. The results of this analysis suggest that the practice effects on retrieval frequency were likely not due to changes in the participants' response biases but to real sensitivity differences (i.e., changes in the memory network).

The increase in retrieval use as a result of practice can be explained by experience-based models (Siegler & Shipley, 1995; Siegler & Shrager, 1984) as well as by the instance theory of automatization (Logan, 1988), whereas the switch from less efficient to more efficient procedural strategies can be accounted for by Siegler's models but not by Logan's model. Note, however, that the frequency of retrieval use never reached 100%. Participants were quite rigid in their strategy choices and continued to use procedural strategies across the experiment. This observation cannot be accounted for by the models discussed above, as these models predict exclusive retrieval use after extensive practice. However, it is also possible that the amount of practice was not extensive enough. Future research might investigate whether more extensive practice would result in exclusive retrieval use.

The reasonably high amount of procedural use after practice is in agreement with Baroody's (1983, 1984, 1985) theory, which states that people prefer procedures above retrieval

because procedural knowledge is cognitively more economical than storing all individual facts in long-term memory. As the storage capacity of long-term memory is unknown, though, this assertion of economy might make no sense. Indeed, in terms of working memory, it is clear that procedural strategy use is *less* economical than retrieval strategy use. Procedural strategies are slower (even after practice) and require more working-memory resources than retrieval strategies do (Imbo & Vandierendonck, in press a,b). Moreover, since transformation strategies often require the retrieval of arithmetic facts, memorization of facts is obligatory for procedural strategy use as well. Baroody's theory is thus unable to explain the replacement of 'pure' procedural strategies (such as counting) by procedural strategies with a retrieval component (such as transformation).

Finally, LeFevre et al. (1996a) also argue that procedural strategies are maintained, even after intensive practice. This can be explained by assuming that the association between a problem and an efficient procedure (e.g., transformation) is as strong as (or even stronger than) the direct association between that problem and its answer. Procedural strategies are then automatically activated, without (or before) the answer is retrieved. This reasoning may also explain why changes in strategy selection are hard to make: when people perceive their commonly used procedural strategy as efficient, why should they switch to retrieval?

Practice effects on strategy efficiency

Practice enhanced strategy efficiency. In Experiment 1, procedural RTs decreased as a result of practice but retrieval RTs did not, whereas in Experiment 2, both retrieval and

procedural RTs decreased as a result of practice. The ASCM (Siegler & Shipley, 1995) and the CMPL theory (Rickard, 1997) predict higher efficiencies in both retrieval and procedural strategies. The instance theory of automatization (Logan, 1988), in contrast, predicts practice effects on retrieval efficiency but not on procedural efficiency, whereas Baroody (1983, 1984, 1985) predicts rather the opposite.

The fact that practice effects differ across strategies, poses problems to the frequently used power law of Newell and Rosenbloom (1981)⁵, which predicts a negatively accelerating rate of speedup as a function of practice. Based on our results, we argue that this law should not be applied to overall latencies, i.e., retrieval and procedural latencies combined. In contrast, we believe that this law should be applied to retrieval and procedural strategies separately. Delaney, Reder, Staszewski, and Ritter (1998) re-analyzed data obtained in previous practice studies with complex-arithmetic problems. They indeed observed that the improvement in latencies was better explained by practice on a strategy than by practice on the task as a whole. This has also been confirmed by Rickard (1997), who showed that the power law does not hold for practice effects on overall latency data, but does hold within each strategy.

Another interesting observation with regard to strategy efficiency was that practice effects on strategy efficiency were larger for multiplication than for addition, and especially when multiplications were solved by procedural strategies. Comparable effects have been observed in a previous study (Imbo, Vandierendonck, & Rosseel, 2007), in which arithmetic experience influenced strategy selection and strategy efficiency for multiplication problems only. There are several possible explanations for this effect. First, addition problems are generally solved faster than multiplication problems (cf. Tables 2 and 4). Hence, increasing the strategy efficiency is easier for multiplication than it is for addition. Second, the procedures used to solve

multiplication problems might be more consistent than those used to solve addition problems. Multiplication problems with a 9 (e.g., 9×7), for example, were consistently solved with the 'ten rule' (e.g., $9 \times 7 = 10 \times 7 - 7 = 70 - 7 = 63$). Once this rule is sufficiently mastered, people are able to use this rule very efficiently (i.e., very fast and accurately). The available rules are less consistent for addition, and most of them involve counting, which is very time-consuming.

The problem-size effect

The present results showed that practice influenced only two out of three sources of the problem-size effect. First, practicing large problems influenced strategy selection, since retrieval was used more often on the practiced (i.e., large) problems as a result of practice. Consequently, the difference in retrieval use between large and small problems became smaller, reducing the problem-size effect. Second, practicing large problems did not change the retrieval problem-size effect. Retrieval was always slower for large problems than for small problems, and the difference in retrieval efficiency between large and small problems was not reduced by practice. Third, practicing large problems did change the procedural problem-size effect. Since procedures became faster for large problems as practice progressed, the procedural problem-size effect decreased. In Experiment 2, the decrease was so strong that the problem-size effect for multiplication inverted. We might thus conclude that more frequent retrieval use on large problems and more efficient procedural use on large problems reduced the problem-size effect as a result of practice. The difference in retrieval efficiency between small and large problems was not influenced by practice.

The retrieval problem-size effect was never eliminated (i.e., large problems were always retrieved more slowly than were small problems), even at asymptotic response times, which is in line with previous research (e.g., Fendrich et al., 1993; Pauli et al., 1998). Yet, several models predict that the problem-size effect should disappear with intensive practice if only retrieval strategies would be used (e.g., Logan, 1988; Siegler & Shipley, 1995; Siegler & Shrager, 1984). The robustness of the retrieval problem-size effect indicates that large problems must have some inherent difficulties in comparison with small problems. There are only a few models that offer candidate explanations for such inherent differences. In Campbell's network-interference model (1995), the presentation of an arithmetic problem primes both the correct answer and its associated (incorrect) answers. Since large problems have more associated answers than small problems, interference is larger for large problems than for small problems. Since this explanation attributes the problem-size effect to structural aspects of memory representations, it should not be eliminated by practice. Similarly, in the interacting neighbors model of Verguts and Fias (2005), the problem-size effect arises because solutions of small problems are more consistent with their neighboring answers than are solutions of large problems.

Can the problem-size effect then ever disappear? Yes, it can, albeit only in extraordinary cases. For example, Geary (1996) observed no problem-size effect in (probably extensively practiced) Chinese children, and Pauli, Lutzenberger, Birbaumer, Rickard, and Bourne (1996) observed no problem-size effect in a mental calculator. Answer magnitude thus produces inherent changes in strategy efficiencies, which can only be overcome, if at all, by very extensive practice.

Transfer effects

Calculation skill learning is very specific, since practice effects were limited to the operation that had been practiced. Practice effects did transfer over complexity though, as practiced participants were better than unpracticed participants in solving complex-arithmetic facts (cf. Experiment 1). It should be noted that transfer only occurred for multiplication but not for addition in Experiment 1, and was (insignificantly) larger for multiplication than for addition in Experiment 2.

The results thus indicate that transfer effects occur rather rarely, and were limited to procedural strategies and absent for retrieval strategies. This might be explained by the fact that procedures are applicable to several problems whereas retrieval is item specific. The data seem to confirm this statement. Indeed, (a) practice enhanced procedural efficiencies more strongly for multiplication than for addition, and (b) transfer occurred more clearly for multiplication than for addition. Stronger transfer effects for procedural strategies than for retrieval have been observed previously (e.g., Campbell, Fuchs-Lacelle, & Phenix, 2006; Delazer et al., 2005).

The observation of transfer effects is in agreement with Baroody's theory (1983, 1984, 1985), which emphasizes the economy and all-round applicability of procedural knowledge. Our data are also in agreement with the identical elements model of Rickard (2005; Rickard et al., 1994), although his theory is more retrieval-based (i.e., suited for simple-arithmetic performance *after* extended practice) than the theories of Baroody (1983, 1984, 1985), which are more procedure-based. The identical elements model assumes a distinct abstract representation for each unique combination of the basic elements (i.e., the operands and the required operation). Transfer is thus possible within the same operation but not between operations, which was observed in the present study. Rickard's (1997) CMPL theory is also consistent with transfer from simple to complex problems, provided that the retrievals involved in simple-arithmetic problem solving

(e.g., $6 + 7$) are also required in complex-arithmetic problem solving (e.g., $16 + 7$). Finally, the instance theory of automatization (Logan, 1988) predicts that transfer to novel stimuli should be poor to nonexistent. As automatization is item-based rather than process-based, it involves learning specific responses to specific stimuli.

Summary

In reviewing several theories and models, we discovered that each model has its strengths and weaknesses. However, no model could explain all practice effects observed. The instance theory of automatization (Logan, 1988) was able to explain the replacement of procedural strategies by direct memory retrieval and the increase in retrieval efficiency. Practice effects on procedural efficiency and the concomitant transfer effects, in contrast, could not be explained by this theory. The latter two effects fit very well in the procedure-based theory of Baroody (1983, 1984, 1985), which is silent about increases in retrieval efficiency and retrieval frequency, though. The ASCM of Siegler and Shipley (1995) and the CMPL theory of Rickard (1997) were able to explain practice effects on strategy selection, retrieval efficiency, and procedural efficiency.

Clearly, we need an integrative model that could account for *all* the practice effects we observed. Instead of developing a new model, we believe that previously developed models can readily be modified in order to account for our data. The ASCM (Siegler & Shipley, 1995), for example, assumes that a particular strategy is selected when that strategy can be executed efficiently. Because retrieval is generally the fastest and most accurate strategy, this strategy will

be selected. Yet, when a nonretrieval strategy can be executed with reasonable speed and accuracy, this strategy also has chances to be selected. The ASCM further predicts that successful strategy execution strengthens problem-answer associations, which results in virtually 100% retrieval use. As we observed that many participants continued to use nonretrieval strategies, we propose a slightly modified version of the ASCM, in which problem-strategy association strengths are maintained relative to problem-answer association strengths. Nonretrieval strategies might then be automatically activated, and, consequently, initiated before retrieval. This slightly modified version of the ASCM would guarantee the continued, if infrequent, use of nonretrieval strategies.

A final question is whether people with strong problem-strategy associations will *ever* switch to retrieval strategies. To ensure progress in people's strategy selection process, the ASCM should incorporate two architectural features that are included in the CMPL theory (Rickard, 1997) as well: (a) nonretrieval strategy execution strengthens retrieval nodes but not vice versa, and (b) retrievals actively inhibit nonretrieval strategies but not vice versa. Differential parameters across participants might then explain why some people are 100% retrieval users and others are not.

References

- Anderson, J.R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.
- Ashcraft, M.H. (1982). The development of mental arithmetic: A chronometric approach. *Developmental Review, 2*, 213-236.
- Ashcraft, M.H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44*, 75-106.
- Baroody, A.J. (1983). The development of procedural knowledge: An alternative explanation for chronometric trends of mental arithmetic. *Developmental Review, 3*, 225-230.
- Baroody, A.J. (1984). A re-examination of mental arithmetic models and data: A reply to Ashcraft. *Developmental Review, 4*, 148-156.
- Baroody, A.J. (1985). Mastery of the basic number combinations: Internalization of relationships or facts? *Journal of Research in Mathematics Education, 16*, 83-98.
- Brigman, S., & Cherry, K.E. (2002). Age and skilled performance: Contributions of working memory and processing speed. *Brain and Cognition, 50*, 242-256.
- Butterworth, B. (1999). *The Mathematical Brain*. London: Macmillan.
- Campbell, J.I.D. (1987). Network interference and mental multiplication. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 13*, 109-123.
- Campbell, J.I.D. (1995). Mechanisms of simple addition and multiplication: A modified network-interference theory and simulation. *Mathematical Cognition, 1*, 121-164.

- Campbell, J.I.D. (1997). On the relation between skilled performance of simple division and multiplication. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *23*, 1140-1159.
- Campbell, J.I.D., & Austin, S. (2002). Effects of response time deadlines on adults' strategy choices for simple addition. *Memory & Cognition*, *30*, 988-994.
- Campbell, J.I.D., & Fuchs-Lacelle, S., & Phenix, T.L. (2006). Identical elements model of arithmetic memory: Extension to addition and subtraction. *Memory & Cognition*, *34*, 633-647.
- 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., Gunter, R. (2002). Calculation, culture, and the repeated operand effect. *Cognition*, *86*, 71-96.
- Campbell, J.I.D., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General*, *130*, 299-315.
- Compton, B.J., & Logan, G.D. (1991). The transition from algorithm to retrieval in memory-based theories of automaticity. *Memory & Cognition*, *19*, 151-158.
- Dehaene, S. (1997). *The Number Sense*. New York: Oxford University Press.
- Delaney, P.F., Reder, L.M., Staszewski, J.J., & Ritter, F.E. (1998). The strategy-specific nature of improvement: The power law applies by strategy within task. *Psychological Science*, *9*, 1-7.

- Delazer, M., Ischebeck, A., Domahs, F., Zamarian, L., Koppelstaetter, F., Siedentopf, C.M., Kaufmann, L., Benke, T., & Felber, S. (2005). Learning by strategies and learning by drill – Evidence from an fMRI study. *Neuroimage*, 25, 838-849.
- Fendrich, D.W., Healy, A.F., & Bourne, L.E. (1993). Mental arithmetic: Training and retention of multiplication skill. In C. Izawa (Ed.), *Applied Cognitive Psychology: Applications of Cognitive Theories and Concepts* (pp. 111-133). Hillsdale, NJ: Erlbaum.
- French, J.W., Ekstrom, R.B., & Price, I.A. (1963). *Kit of reference tests for cognitive factors*. Princeton, NJ: Educational Testing Service.
- Fuson, K.C. (1982). An analysis of the counting-on procedure in addition. In T.H. Carpenter, J.M. Moser, & T.H. Romberg (Eds), *Addition and subtraction: A cognitive perspective* (pp. 67-78). Hillsdale NJ: Lawrence Erlbaum.
- Fuson, K.C. (1988). *Children's counting and concepts of number*. New York: Springer.
- Geary, D.C. (1996). The problem-size effect in mental addition: Developmental and cross-national trends. *Mathematical Cognition*, 1996, 63-93.
- Geary, D.C., Hamson, C.O., Chen, G.-P., Fan, L., Hoard, M.K., & Salthouse, T.A. (1997). Computational and reasoning in arithmetic: Cross-generational change in China and the United States. *Psychonomic Bulletin and Review*, 4, 254-263.
- Geary, D.C., Salthouse, T.A., Chen, G.-P., & Fan, L. (1996). Are East Asian versus American differences in arithmetic ability a recent phenomenon? *Developmental Psychology*, 32, 254-262.
- Gelman, R., & Gallistel, R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

- Hoyer, W.J., Cerella, J., & Onyper, S.V. (2003). Item learning in cognitive skill training: Effects of item difficulty. *Memory & Cognition*, *31*, 1260-1270.
- Imbo, I., & Vandierendonck, A. (in press a). The role of phonological and executive working-memory resources in simple-arithmetic strategies. *European Journal of Cognitive Psychology*.
- Imbo, I., & Vandierendonck, A. (in press b). Do multiplication and division strategies rely on executive and phonological working-memory resources? *Memory & Cognition*.
- 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.
- 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.
- Klapp, S.T., Boches, C.A., Trabert, M.L., Logan, G.D. (1991). Automatizing alphabet arithmetic: II. Are there practice effects after automaticity is achieved? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *17*, 196-209.
- Kroesbergen, E.H., & Van Luit, J.E.H. (2003). Mathematics interventions for children with special educational needs: A meta-analysis. *Remedial and Special Education*, *24*, 97-114.
- LeFevre, J.-A., Bisanz, J., Daley, K.E., Buffone, L., Greenham, S.L., & Sadesky, G.S. (1996a). Multiple routes to solution of single-digit multiplication problems. *Journal of Experimental Psychology: General*, *125*, 284-306.
- LeFevre, J.-A., & Liu, J. (1997). The role of experience in numerical skill: Multiplication

- performance in adults from Canada and China. *Mathematical Cognition*, 3, 31-62.
- LeFevre, J.-A., & Morris, J. (1999). More on the relation between division and multiplication in simple arithmetic: Evidence for mediation of division solutions via multiplication. *Memory & Cognition*, 27, 803-812.
- LeFevre, J.-A., Sadesky, G.S., & Bisanz, J. (1996b). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 22, 216-230.
- Lewis, A.B. (1989). Training students to represent arithmetic word problems. *Journal of educational psychology*, 81, 521-531.
- Logan, G.D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95, 492-527.
- Logan, G.D., & Klapp, S.T. (1991). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 17, 179-195.
- McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Brain and Cognition*, 4, 171-196.
- Mulhern, G., & Wylie, J. (2004). Changing levels of numeracy and other core mathematical skills among psychology undergraduates between 1992 and 2002. *British Journal of Psychology*, 95, 355-370.
- Newell, A., & Rosenbloom, P.S. (1981). *Mechanisms of skill acquisition and the law of practice*. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.

- Onyper, S.V., Hoyer, W.J., & Cerella, J. (2006). Determinants of retrieval solutions during cognitive skill training: Source confusions. *Memory & Cognition, 34*, 538-549.
- Pauli, P., Bourne, L.E., & Birbaumer, N. (1998). Extensive practice in mental arithmetic and practice transfer over a ten-month retention interval. *Mathematical Cognition, 4*, 21-46.
- Pauli, P., Lutzenberger, W., Birbaumer, N., Rickard, T.C., & Bourne, L.E. (1996). Neurophysiological correlates of mental arithmetic. *Psychophysiology, 33*, 522-529.
- Pauli, P., Lutzenberger, W., Rau, H., Birbaumer, N., Rickard, T.C., Yaroush, R.A., & Bourne, L.E. (1994). Brain potentials during mental arithmetic – Effects of extensive practice and problem difficulty. *Cognitive Brain Research, 2*, 21-29.
- Rickard, T.C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General, 126*, 288-311.
- Rickard, T.C. (2004). Strategy execution in cognitive skill learning: An item-level test of candidate models. *Journal of Experimental Psychology, 30*, 65-82.
- Rickard, T.C. (2005). A revised identical elements model of arithmetic fact representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 250-257.
- Rickard, T.C., & Bourne, L.E. (1996). Some tests of an identical elements model of basic arithmetic skills. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1281-1295.
- Rickard, T.C., Healy, A.F., & Bourne, L.E. (1994). On the cognitive structure of basic arithmetic skills: Operation, order, and symbol transfer effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20*, 1139-1153.

- Royer, J.M., Tronsky, L.N., Chan, Y., Jackson, S.J., & Marchant, H. III. (1999). Math-fact retrieval as the cognitive mechanism underlying gender differences in math test performance. *Contemporary Educational Psychology, 24*, 181-266.
- Seyler, D.J., Kirk, E.P., & Ashcraft, M.H. (2003). Elementary subtraction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1339-1352.
- Shrager, J., & Siegler, R.S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science, 9*, 405-410.
- Siegler, R.S. (1988). Strategy choice procedures and the development of multiplication skill. *Journal of Experimental Psychology: General, 117*, 258-275.
- Siegler, R.S., & Araya, R. (2005). A computational model of conscious and unconscious strategy discovery. *Advances in Child Development and Behavior, 33*, 1-42.
- Siegler, R.S., & Shipley, C. (1995). Variation, selection, and cognitive change. In G. Halford & T. Simon (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31-76). Hillsdale, NJ: Erlbaum.
- Siegler, R.S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229-294). Hillsdale, NJ: Erlbaum.
- Spelke, E., & Dehaene, S. (1999). Biological foundations of numerical thinking: response to T.J. Simon. *Trends in Cognitive Science, 3*, 365-366.
- Spelke, E.S., & Tsivkin, S. (2001). Language and number: a bilingual training study. *Cognition, 78*, 45-88.

- Steel, S., & Funnell, E. (2001). Learning multiplication facts: A study of children taught by discovery methods in England. *Journal of Experimental Child Psychology, 79*, 37-55.
- Touron, D.R., Hoyer, W.J., & Cerella, J. (2004). Cognitive skill learning : Age-related differences in strategy shifts and speed of component operations. *Psychology and Aging, 19*, 565-580.
- Van der Goten, K., & Vandierendonck, A. (1997). Software to calculate measures of sensitivity and response bias based on detection theory and threshold theory. *Behavior Research Methods, Instruments, & Computers, 29*, 461-463.
- Verguts, T., & Fias, W. (2005). Interacting neighbors: A connectionist model of retrieval in single-digit multiplication. *Memory & Cognition, 33*, 1-16.
- Whetstone, T. (1998). The representation of arithmetic facts in memory: Results from retraining a brain-damaged patient. *Brain and Cognition, 36*, 290-309.
- Wynn, K. (1990). Children's understanding of counting. *Cognition, 36*, 155-193.
- Zbrodoff, N.J. (1999). Effects of counting in alphabet arithmetic: Opportunistic stopping and priming of intermediate steps. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 299-317.

Author note

This research reported in this article was supported by grant no. 011D07803 of the Special Research Fund at Ghent University to the first author. Correspondence concerning this article should be addressed to Ineke Imbo, Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 B-Ghent, Belgium. E-mails may be sent to Ineke.Imbo@UGent.be

Footnotes

1. In the present paper, the term ‘training’ is seen as the explicit training of a particular strategy, whereas the term ‘practice’ is seen as exercising through repetition, without enforcing a particular strategy.
2. There exist several recent adaptations of the ASCM, such as the strategy choice and discovery simulation (SCADS) model (Shrager & Siegler, 1998), which incorporates the ASCM but also models metacognitive processes to allow for the discovery of new strategies; and the SCADS* model (Siegler & Araya, 2005), which adds six new mechanisms to the SCADS model (i.e., controlled attention, interruption of procedures, verbalization, priming, forgetting, and dynamic feature detection).
3. Note that the design used in this first experiment (i.e., without a pretest but with different practice groups) has been used earlier (e.g., Rickard et al., 1994; Rickard & Bourne, 1996). We acknowledge that the absence of a pretest is a drawback of the first experiment. Therefore, a pretest has been included in the second experiment.

4. As the amount of counting strategy use is very low (cf. Tables 1 and 3), we decided to put counting RTs and transformation RTs together. These procedural RTs are then contrasted with the retrieval RTs. Though, since (a) not all strategies were used across all the practice sessions, and (b) only RTs of the correctly solved problems were analyzed, for some subjects empty cells occurred in the practice level x operation x size x strategy ANOVA. We replaced these empty cells for each participant with the correct RT of the corresponding cell [i.e., the mean RT (over participants) of the practice level x operation x size x strategy cell]. Obviously, this procedure was only needed in the ANOVAs on strategy efficiency and not in the ANOVAs on strategy selection. The number of cells replaced was 10 (out of 240) in Experiment 1 and 34 (out of 320) in Experiment 2.

5. $T = BN^{-\alpha}$ in which T represents the performance time, B is the time taken to perform the first trial, N is the trial number, and α represents the rate at which performance time changes.

Table 1

Percentages of retrieval, transformation, and counting use as a function of problem size, operation, and practice level (Experiment 1). Standard errors are shown between brackets.

		Addition		Multiplication	
		Small	Large	Small	Large
Retrieval	0x practice	72 (7)	46 (9)	91 (7)	56 (9)
	3x practice	56 (4)	39 (9)	82 (7)	69 (9)
	6x practice	70 (7)	58 (9)	87 (7)	90 (10)
Transformation	0x practice	24 (4)	50 (9)	3 (4)	21 (9)
	3x practice	21 (4)	54 (9)	3 (4)	25 (9)
	6x practice	24 (4)	42 (9)	3 (4)	9 (9)
Counting	0x practice	3 (5)	3 (4)	6 (5)	23 (4)
	3x practice	23 (5)	6 (4)	15 (5)	6 (4)
	6x practice	6 (5)	0 (4)	9 (5)	0 (4)

Table 2

Retrieval and procedural response times (in msec) as a function of problem size, operation, and practice level (Experiment 1). Standard errors are shown between brackets.

		Addition		Multiplication	
		Small	Large	Small	Large
Retrieval	0x practice	889 (59)	986 (103)	1041 (59)	1300 (103)
	3x practice	795 (59)	957 (103)	994 (59)	1212 (103)
	6x practice	838 (59)	923 (103)	1025 (62)	1296 (109)
Procedural	0x practice	1282 (327)	1303 (260)	3666 (327)	4213 (260)
	3x practice	1146 (327)	1115 (260)	2272 (327)	1908 (260)
	6x practice	1418 (327)	1311 (260)	2300 (344)	2457 (274)

Table 3

Percentages of retrieval, transformation, and counting use as a function of problem size, operation, and session (Experiment 2). Standard errors are shown between brackets.

		Addition		Multiplication	
		Small	Large	Small	Large
Retrieval	Pretest	77 (3)	46 (7)	94 (3)	69 (7)
	Posttest	74 (3)	57 (6)	94 (3)	84 (6)
Transformation	Pretest	20 (2)	52 (7)	4 (2)	30 (7)
	Posttest	20 (2)	43 (6)	2 (2)	16 (6)
Counting	Pretest	4 (1)	2 (1)	2 (1)	1 (1)
	Posttest	6 (2)	0 (0)	3 (2)	0 (0)

Table 4

Retrieval and procedural response times (in msec) as a function of problem size, operation, and session (Experiment 2). Standard errors are shown between brackets.

		Addition		Multiplication	
		Small	Large	Small	Large
Retrieval	Pretest	741 (45)	904 (93)	941 (45)	1222 (93)
	Posttest	706 (37)	847 (65)	889 (37)	1084 (65)
Procedural	Pretest	1193 (137)	1182 (119)	1823 (137)	1722 (119)
	Posttest	1156 (146)	1104 (83)	1941 (146)	1606 (83)

Table 5

Values of sensitivity (d') and response bias (c) as a function of session and operation (Experiment 2). Standard errors are shown between brackets.

	Sensitivity (d')		Response bias (c)	
	Addition	Multiplication	Addition	Multiplication
Pre-practice	2.1 (0.2)	3.1 (0.2)	-0.1 (0.1)	-0.1 (0.1)
Post-practice	2.1 (0.2)	3.4 (0.2)	-0.1 (0.1)	0.0 (0.1)