Instruction and load effects on high-skill and low-skill individuals:

A study in the domain of mental arithmetic

Ineke Imbo & André Vandierendonck

Ghent University, Belgium

Running head: Instruction and load effects in mental 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

What happens when people are asked to respond as quickly or as accurately as possible? This study tested the effects of speed/accuracy instructions and working-memory load on people's strategy efficiency and strategy selection. Adult participants solved simple addition problems (Experiment 1) and simple multiplication problems (Experiment 2) under load and no-load conditions and provided trial-by-trial strategy reports. High-skill participants were more efficient than low-skill participants, but the underlying causes of these skill-related effects differed across experiments. In the addition experiment, high-skill participants responded adaptively to the changing situations by changing their strategy choices, which resulted in smaller effects on their actual performance. Low-skill participants in contrast, did not change their strategy choices as adaptively, which resulted in less efficient performance – and especially so under load conditions. In the multiplication experiment, high-skill and low-skill participants differed in strategy efficiencies rather than in strategy choices. In the discussion, the results are further interpreted and future adaptations for the adaptive strategy choice model (ASCM, Siegler & Jenkins, 1989) are suggested.

Acknowledgements

Support for this research was provided by the Research Foundation Flanders (FWO Flanders) with a postdoctoral fellowship to Ineke Imbo. Thanks are extended to the reviewers for their helpful comments on previous drafts of this paper.

Instruction and load effects on high-skill and low-skill individuals: A study in the domain of mental arithmetic

The world is an ever-changing place. Even if we repeatedly perform the same task, several characteristics can urge us to change our behavior. For example, in some situations you need a quick answer but it doesn't need to be highly precise, whereas in other situations you don't necessarily need to be fast, as long as the answer is 100% correct. Similarly, in some situations you have all your working memory resources available, whereas in other situations a part of your working memory is loaded by another task or by stress. The question is now: how do people react to such situational changes? Will they choose other strategies and so mitigate performance decrements (in terms of speed and accuracy)? Or will they rather continue using the same strategies in all the different situations, and thus show great changes in their actual performance (e.g., respond faster but less accurately under speed instructions and more accurately but slower under accuracy instructions)? We expect that the answer to this question will differ as a function of skill. More specifically, we suppose that high-skill participants will adaptively change their behavior according to the situation, whereas low-skill participants will rather undergo the changing situations. As a result, we expect that the effects of instruction and working-memory load will be greater for low-skill than for high-skill participants.

In the present study, we tested these predictions in the domain of mental arithmetic. In daily life, we often have to solve arithmetic problems such as 9 + 7 and 4 x 8; and most adults are able to solve these problems with a reasonable speed and accuracy. However, as in many other cognitive processes, in some situations speed is more important than accuracy and in other situations accuracy is more important than speed. According to the Adaptive Decision Maker of Payne, Bettman, and Johnson (1993), people adapt to changing situations in three ways: by speeding up/down the execution of certain strategies, by applying the strategies in a more precise/imprecise way, and by selecting other strategies. Applied to the domain of mental arithmetic, this implies that speed and accuracy instructions can change participants' response times (e.g., faster responses under speed instructions than under accuracy instructions), participants' accuracies (e.g., higher accuracies under accuracy instructions than under speed instructions), and participants' strategy choices (e.g., more frequent retrieval use under speed instructions than under accuracy instructions). The decision to test our predictions in the domain of mental arithmetic also allowed us to frame our research questions and conclusions in a theoretical model that is specific for the mathematical domain, namely the adaptive strategy choice model (ASCM, Siegler & Jenkins, 1989; Siegler & Shipley, 1995).

The distribution of associations model of Siegler and Shrager (1984) is the precursor of the ASCM, and will be introduced first. In the distribution of associations model, the basic representation of a problem is accompanied not only by its correct answer (e.g., 6 x 7 = 42), but also by incorrect answers that have been generated across experience (e.g., 6 x 7 = 49). Problems have peaked distributions when the association between the problem and its correct answer is strong while the associations with other answers are weak. Problems have flat distributions of associations when the association with the correct answer is only slightly stronger than the associations with other, incorrect answers. The more peaked the distribution of associations, the higher the probability of retrieval. Due to past experiences, small problems have more peaked distributions of associations than do large problems, resulting in the well-known problem size effect.

In the later developed ASCM, people do not only accumulate information about answers to arithmetic problems, but also about strategy efficiencies (i.e., speed and accuracy). This strategy information then determines the problem-strategy association strength. The model thus first selects a strategy based on the distribution of strategy strengths, and then attempts to execute that strategy. Whether one first attempts to solve the problem with a retrieval or non-retrieval strategy (such as transformation or counting) depends on the relative association strengths. Problems with flat distributions of associations generally have weak problem-answer associations but strong associations between the problem and non-retrieval strategies. Accordingly, the probability of retrieving an answer on the first retrieval attempt is small; if the answer is retrieved at all, multiple retrieval attempts will be required. If the answer cannot be retrieved, a non-retrieval strategy will be used to solve the problem. Problems with peaked distributions of associations, in contrast, will readily be solved with the retrieval strategy.

One will thus retrieve an answer from long-term memory only if the problem can be solved fast and accurately with the retrieval strategy. Stated differently, retrieval is applied when the association strength does exceed the confidence criterion (which determines how sure one must be to state a retrieved answer) and when the search length criterion (which determines how many attempts one will make to retrieve an answer before trying a non-retrieval strategy) is not exceeded. If the retrieval strategy would provide a slow and/or incorrect answer (e.g., when a problem is associated with several possible answers in long-term memory) and thus exceeds the search length criterion or does not exceed the confidence criterion, a non-retrieval strategy will be used. The interesting fact about the ASCM is that people can change their confidence criterion and search length criterion according to the situation. For example, more frequent retrieval use can be obtained by applying a less strict confidence criterion or by applying a longer search length criterion. As will be explained below, we predict that the instructions imposed by the experimenter will encourage participants to change their criteria and – consequently – their strategy choices. However, it is also possible that participants are reluctant to change criteria. In that case, they might continue to use slow and resource-demanding non-retrieval strategies instead of faster and less resource-demanding retrieval strategies, which results in less efficient arithmetic performance.

Although arithmetic performance is influenced by various individual differences such as arithmetic skill, calculator use, gender, math anxiety and mathematical experience (e.g., Imbo, Vandierendonck, & Rosseel, 2007; Imbo & Vandierendonck, 2007b,c), we decided to focus on one specific individual difference, namely arithmetic skill. In terms of the ASCM, high-skill participants have more peaked distributions of associations and stronger problem-answer associations than low-skill participants. Accordingly, high-skill participants would use the retrieval strategy more frequently and would respond faster and more accurately. However, it is also possible that high-skill participants use a stricter confidence criterion, which would mean that they use the retrieval strategy almost as frequently as low-skill participants, but still respond faster and more accurately. This would also imply that high-skill participants have more space to adjust their confidence criterion to changing situations (e.g., speed instructions or load conditions).

We will deal with two questions. First, do high- and low-skill participants react differently on speed and accuracy instructions? And second, are high- and low-skill participants differently affected by a working-memory load? Smith-Chant and LeFevre (2003) already showed that low-skill participants are more affected by speed and accuracy demands than high-skill participants. We extended Smith-Chant and LeFevre's study by not only including 'pure' speed and accuracy instructions, but also a speed/accuracy instruction in which people have to respond fast *and* accurately. We expect larger trade-offs between speed and accuracy in low-skill than in high-skill participants. We further predict that high-skill participants will respond more adaptively to the changing instructions than will low-skill participants. As noted above, adaptive behavior can be obtained by changing strategy efficiencies and by changing strategy choices (e.g., adjust the confidence criterion in order to increase the use of the fast retrieval strategy).

We also tested the working-memory requirements in high-skill and low-skill participants. Working memory can be defined as a set of processing resources of limited capacity, involved in information maintenance and processing (e.g., Baddeley & Logie, 1999; Engle, Tuholski, Lauglin, & Conway, 1999; Miyake, 2001) and has been shown to play a significant role in adults' simple-arithmetic performance (see DeStefano & LeFevre, 2004, for review). Because low-skill participants use more non-retrieval strategies than do high-skill participants, our prediction concerning strategy efficiency is that low-skill people will need more working memory resources than high-skill people. Concerning strategy selection, no load effects were expected. Previous studies already showed that, although being fast and accurate loads on working-memory resources, choosing between strategies does not (e.g., Hecht, 2002, Imbo & Vandierendonck, 2007a,b,c).

Because there exist strategic differences between addition and multiplication (i.e., multiplication is more retrieval-based than addition; Campbell & Xue, 2001; Hecht, 1999; Imbo, Vandierendonck, & Rosseel, 2007), the effects of instruction and working-memory load were tested in two experiments – one for each operation. As suggested by Imbo and Vandierendonck (2008), the difference between addition and multiplication can be framed within the ASCM by supposing that the distributions of associations are more highly peaked for addition than for multiplication, in combination with lower confidence criteria and/or longer search lengths for multiplication than for addition. The different memory structure for addition and multiplication also exerts different predictions for each experiment. Because the distributions of associations are so highly peaked for addition, we predict that it should be easy to increase retrieval use (e.g., by lowering the confidence criterion) without losing speed or accuracy. Because the distributions of associations are less highly peaked for multiplication, people will not easily switch strategies in the multiplication experiment. We predict that, when confronted with speed or accuracy instructions, participants in the multiplication experiment will change the efficiency of the retrieval strategy itself rather than switching to another strategy. Another reason why multiplication should be more resistant to strategy shifts than addition is the difference in non-retrieval efficiencies. For multiplication most non-retrieval strategies are highly inefficient, whereas there exist many efficient non-retrieval strategies for addition. We further predict that changes in strategy selection (for addition) and strategy efficiency (for multiplication) will be significant for high-skill participants only. Because low-skill participants may already perform as good as they can (in terms of both strategy selection and strategy efficiency), they may have no room to change their performance and will not be able to respond to the instructions in a similar way.

Experiment 1: Addition

Method

Participants. Forty subjects participated in the present experiment (21 men and 19 women; mean age 19 years 3 months). All subjects were first-year psychology students at Ghent University and participated for course requirements and credits.

The addition task. The addition problems consisted of two one-digit numbers (e.g., 6 + 7). Problems involving 0 or 1 as an operand (e.g., 5 + 0, 1 + 4) and tie problems (e.g., 3 + 3) were excluded. Since commuted pairs (e.g., 9 + 4 and 4 + 9) were considered as two different problems, this resulted in 56 addition problems (ranging from 2 + 3 to 9 + 8). If the sum of both operands was smaller than 10, the problem was defined as small. If the sum of both operands was larger than 10, the problem was defined as large.

The executive secondary task. A continuous choice reaction time task (CRT task) was used to load the executive working-memory component. Szmalec, Vandierendonck, and Kemps (2005) have shown that this task interferes with the central executive while the load on the phonological and visuo-spatial slave systems is negligible. Stimuli of the CRT task consisted of low tones (262 Hz) and high tones (524 Hz) that were sequentially presented with an interval of 900 or 1500 ms. Participants had to press the 4 on the numerical keyboard when they heard a high tone and the 1 when a low tone was presented. The duration of each tone was 200 ms. The tones were presented continuously during the simple-arithmetic task. The CRT task was also performed alone in two control conditions (i.e., without the concurrent solving of arithmetic problems). In the first one, the CRT task was performed without any concurrent task. In the second one, the arithmetic problems and their correct answer were presented, which the participants had to read off the screen. This allowed us to test whether the secondary task was affected by the central calculation processing or by other processes (such as encoding and responding).

Procedure. Each participant was tested individually in a quiet room for approximately 1 hour. The experiment started with short questions about the age of the participant, calculator use (on a rating scale from 1 “never” to 5 “always”), mathematical experience (i.e., the amount of mathematical lessons in the last year of high school), and math anxiety (on a rating scale from 1 “low” to 5 “high”). These individual difference variables were used to test whether high- and low skill differed on other aspects than arithmetic skill[1]. All participants solved the simple-arithmetic problems in three conditions: (1) a speed condition, in which they were asked to solve the problems as fast as possible, (2) an accuracy condition, in which they were asked to solve the problems as accurately as possible, and (3) an accuracy/speed condition, in which they were asked to respond as fast and as accurately as possible. These instruction conditions were randomized across participants. Each condition was further divided in two blocks: one without working-memory load and one in which the executive working-memory component was loaded. For half of the participants, each condition started with the no-load block and was followed by the working-memory load block; the order was reversed for the other half of the participants. All 56 problems were presented once in each Instruction x Load condition.

A trial started with a fixation point for 500 milliseconds. Then the arithmetic problem was presented horizontally in the center of the screen, with the operation sign at the fixation point. The problem remained on screen until the subject responded. Timing began when the stimulus appeared and ended when the response triggered the sound-activated relay. To enable this sound-activated relay, participants wore a microphone that was activated when they spoke their answer aloud. This microphone was connected to a software clock accurate to 1 ms. On each trial, feedback was presented to the participants. In the accuracy condition, a green ‘Correct’ was presented when the answer was correct, and a red ‘Incorrect’ when it was not. In the speed condition, the RT was presented. In the accuracy/speed condition, a combination of both feedback types was presented. Additionally, after each series of 10 trials, the participants’ performance was summarized. That is, in the speed condition, the participant’s mean RT was presented; in the accuracy condition, the participant’s total number of errors was presented; and a combination of both accuracy and speed information was presented in the accuracy/speed condition.

Immediately after solving each problem, participants were presented three types of strategies on the screen: Retrieval, Procedural, and Other. These three types had been extensively explained by the experimenter: (1) Retrieval: You solve the problem by remembering or knowing the answer directly from memory. It means that you know the answer without any additional processing. For example: you know that 5 + 6 = 11 because 11 “pops into your head”. (2) Procedural: It means that you don’t know the answer without any additional processing. One possibility is that you solve the problem by counting one-by-one in order to get the answer. For example: 7 + 4 = 7… 8… 9… 10 … 11. Another possibility is that you solve the problem by making an intermediate step. For example: 8 + 9 = 8 + 2 + 7 = 10 + 7 = 17 or 6 + 7 = 6 + 6 + 1 = 12 + 1 = 13. (3) Other: You solve the problem by a strategy unlisted here, or you do not know what strategy you used to solve the problem. For example: guessing. After each problem, participants were asked to report verbally which of these strategies they had used. The experimenter also emphasized that the presented strategies were not meant to encourage use of a particular strategy. If the participant felt like using only one of the presented strategy types, he/she was completely free to do so; when the participant acknowledged generally using a mix of strategy types; he/she was as free to do so[2]. The answer of the participant, the strategy information, and the validity of the trial were recorded on-line by the experimenter. All invalid trials (e.g., failures of the voice-activated relay) were discarded and returned at the end of the block, which minimized data-loss due to unwanted failures.

The French Kit. After the arithmetic experiment, all participants completed a paper-and-pencil test of arithmetic skill, the French Kit (French, Ekstrom, & Price, 1963). The test consisted of two subtests, one page with complex addition problems and one page with complex subtraction and multiplication problems. Participants were given 2 minutes per page, and were instructed to solve the problems as fast and as accurately as possible. The amount of correct answers on both subtests was summed to yield a total score. In order to test the effect of arithmetic skill, the group of participants was divided into two subgroups, based on the median score (26). The low-skill participants’ score was below or equal to the median score (N = 21;

mean = 22; range = 16 to 26), whereas the high-skill participants' score was above the median score (*N* = 19; mean = 32, range = 27 to 53).

Results and discussion

Of all trials, 12% was spoiled due to failures of the sound-activated relay. Since all these invalid trials returned at the end of the block, most of them were recovered from data loss, which reduced the trials due to failures of the sound-activated relay to 2%. All trials on which participants reported having used a strategy 'Other' (0.5%) were deleted as well. The results section contains four parts. We first tested whether the secondary task exerted its effects as predicted. Next, we analyzed the effects of problem size, instruction, load, and skill on the three main components of arithmetic performance: speed, accuracy, and strategy use. All ANOVA analyses were based on the multivariate general linear model, and all reported results were significant at $p < .05$, unless mentioned otherwise.

Secondary task performance. As noted above, performance on the CRT task was not only obtained during the three different instruction conditions (accuracy, accuracy/speed, speed) but also in two control conditions, one without primary task and one in which participants had to read the arithmetic answers of the screen. An ANOVA was run on the CRT task's RTs (of correctly answered trials only); with Skill as between-subjects factor and repeated measures on Condition (single, naming, accuracy, accuracy/speed, speed; see Table 1). The significant effect of Condition, $F(4,35) = 18.91$, $MSe = 4248$, indicated that participants responded faster in the single task condition (575 ms) than in the condition in which they only had to read the answers off the screen (679 ms), $F(1,38) = 40.38$. RTs in the latter condition were faster than those observed under speed (698 ms), accuracy/speed (690 ms), or accuracy instructions (710 ms), $F(1,38) =$

4.93. Participants responded equally fast in the latter three conditions; and the effect of Skill and the Skill x Condition interaction did not reach significance.

The observation that performance on the CRT task was worse under dual-task conditions than under the single-secondary-task condition indicates that the arithmetic task and the CRT task load common working memory resources. Therefore, load effects on the arithmetic task cannot be due to a tradeoff between the primary task and the secondary task. We also observed that the processes of encoding and responding required a reasonable amount of working memory resources. The process of retrieving or calculating the correct answer on the arithmetic task did still require more working memory resources, though. There was no main effect of skill and no interaction between skill and condition, which again indicates that the primary task results can be safely interpreted.

Insert Table 1 about here

Primary task performance: speed. A 2 x 3 x 2 x 2 ANOVA was run on RTs (of correctly solved problems only, see Table 2) with Arithmetic skill (low vs. high) as between-subjects factor and Instruction (accuracy vs. speed vs. accuracy/speed), Load (no-load vs. load), and Problem size (small vs. large) as within-subject factors. The main effect of Instruction, $F(2,37) = 16.95$, $MSe = 67658$, indicated that the addition performance was faster under speed instructions (851 ms) than under accuracy/speed instructions (908ms), $F(1,38) = 6.88$, and faster under accuracy/speed instructions than under accuracy instructions (1047ms), $F(1,38) = 19.84$. Participants were also faster on small problems (812 ms) than on large problems (1059 ms), $F(1,38) = 152.94$, $MSe = 47559$, and faster in no-load conditions (836 ms) than under working-memory load (1034ms), $F(1,38) = 98.55$, $MSe = 47740$. The main effect of Skill indicated that high-skill participants (862 ms) were faster than low-skill participants (1009 ms), $F(1,38) = 5.65$, $MSe = 460774$.

Skill interacted with Instruction, $F(2,37) = 3.33$, $MSe = 67658$, and with Load, $F(1,38) = 4.55$, $MSe = 47740$. The Skill x Instruction interaction indicated that low-skill participants were more influenced by instructions than were high-skill participants. More specifically, the difference in RTs between accuracy and speed instructions was much larger in low-skill participants (281 ms) than in high-skill participants (111 ms), $F(1,38) = 6.53$. Low-skill participants decelerated tremendously when they were asked to focus on accuracy. The Skill x Load interaction indicated that the load effects were higher for low-skill participants (240 ms) than for high-skill participants (155 ms). Both two-way interactions were further modified by the three-way interaction between Skill, Instruction, and Load, $F(2,37) = 3.14$, $MSe = 36584$. As can be seen in Figure 1, the deceleration in low-skill participants under accuracy instructions was boosted under load conditions.

In order to test whether the aforementioned differences between high- and low-skill participants were due to changed strategy efficiencies, a similar 2 x 3 x 2 x 2 ANOVA was run on pure *retrieval* RTs (of correctly solved problems only; see Table 3). All main effects were significant, $F(2,37) = 12.5$, $MSe = 60698$ for Instruction, $F(1,38) = 110.32$, $MSe = 17487$ for Problem size, $F(1,38) = 71.88$, $MSe = 63545$ for Load, and $F(1,38) = 4.68$, $MSe = 331518$ for Skill. However, Skill did not interact with Instruction, $F(2,37) = 2.52$, $MSe = 60698$ ($p = .10$) or Load, $F(1,38) = 2.73$, $MSe = 63545$ ($p = .11$). That is, the retrieval speed of high- and low-skill participants was *not* differently affected by instruction or load.

Primary task performance: errors. A similar 2 x 3 x 2 x 2 ANOVA was run on error percentages (see Table 2). The main effect of Instruction, $F(2,37) = 13.89$, $MSe = 16$, indicated that people made more errors in the speed condition (4.8%) than in the accuracy/speed condition

(2.9%), $F(1,38) = 18.12$, and more errors in the accuracy/speed condition than in the accuracy condition (1.9%), $F(1,38) = 7.40$. Participants were also more erroneous on large problems (4.9%) than on small problems (1.6%), $F(1,38) = 22.74$, $MSe = 57$, and more erroneous in load conditions (3.5%) than in no-load conditions (2.9%), $F(1,38) = 3.63$, $MSe = 14$ ($p = .06$). The main effect of Skill indicated that low-skill participants (4.2%) were more erroneous than high-skill participants (2.2%), $F(1,38) = 3.98$, $MSe = 117$. Skill did not interact with any variable (each $p > .10$).

The same 2 x 3 x 2 x 2 ANOVA was run on error percentages of *retrieval* trials only (see Table 3). The main effects of Instruction and Size were significant, $F(2,37) = 14.21$, $MSe = 19$, and $F(1,38) = 19.26$, $MSe = 31$. The main effect of Load did not reach significance though, $F<1$. The main effect of Skill was not significant either, $F(1,38) = 2.31$, $MSe = 44$ ($p = .14$), and Skill did not interact with Instruction, $F(2,37) = 2.27$, $MSe = 19$ ($p = .12$) or Load ($F<1$). Instruction effects and load effects were thus equally large in high-skill and low-skill participants' error rates for retrieval-only trials.

Primary task performance: strategy choices. In order to test whether the aforementioned effects on accuracy and speed were due to changes in participants' strategy choices, a similar 2 x 3 x 2 x 2 ANOVA was run on percentages retrieval use (of correctly solved problems only; see Table 2). Participants chose the retrieval strategy more frequently under speed instructions (79.7%) than under accuracy instructions (77.0%), $F(1,38) = 4.66$, $MSe = 108$, and more frequently on small problems (93.6%) than on large problems (63.1%), $F(1,38) = 105.23$, $MSe = 1053$. As expected, participants chose the retrieval strategy as frequently in load and no-load conditions, $F(1,38) = 1.52$ ($p = .23$). The effect of Skill did not reach significance either, $F<1$.

Skill interacted with Load, $F(1,38) = 6.94$, $MSe = 61$. Working memory load did not affect the amount of retrieval use in low-skill participants (range 76.9% - 77.9%), but high-skill participants chose the retrieval strategy more frequently in load conditions (80.7%) than in no-load conditions (77.9%), $F(1,38) = 7.35$. As will be explained in the discussion, this interaction indicates that high-skill participants make more adaptive strategy choices than do low-skill participants. This two-way interaction was further modified by the three-way interaction between Skill, Load, and Size, $F(1,38) = 5.94$, $MSe = 51$. As can be seen in Figure 2, high-skill participants increased their retrieval use under load conditions especially for large problems, which strengthens our conclusion that high-skill participants are more adaptive than low-skill participants.

Insert Tables 2 and 3 and Figures 1 and 2 about here

Discussion. Instruction had an effect on participants' RTs, error rates and strategy choices. As predicted, participants were more accurate under accuracy instructions and faster under speed instructions. Interestingly, the performance under speed/accuracy instructions was always in between the performances observed in speed and accuracy instructions. This indicates that there is always a tradeoff between speed and accuracy: Participants who focus on speed will lose accuracy, and the other way around. The changes in accuracy and speed were accompanied by strategy selection effects. People switched to the faster retrieval strategy under speed conditions, but reverted to slower non-retrieval strategies under accuracy instructions. These adaptive strategy switches can be accounted for by the ASCM, as will be outlined in the general discussion.

High-skill participants were faster and more accurate than low-skill participants. Furthermore, the RTs of high-skill participants were more stable than those of low-skill participants. That is, the slow down under accuracy instructions was much larger for low-skill than for high-skill participants (see also Smith-Chant & LeFevre, 2003). We also observed larger load effects on RTs in low-skill than in high-skill participants. The analyses on retrieval speed and retrieval error rates showed that high- and low-skill participants were not differently affected by instruction or load. This indicates that the instruction and load effects on high- and low-skill participants were caused by changes in strategy selection rather than by changes in strategy efficiency. Indeed, high-skill participants, but not low-skill participants, switched to more frequent retrieval use under load conditions – and especially so for large problems. Since retrieval strategies are faster (e.g., LeFevre et al., 1996a,b) and less resource-demanding (e.g., Imbo & Vandierendonck, 2007a,b,c) than non-retrieval strategies, effects of working memory load are smaller in high-skill than in low-skill participants. The adaptive strategy switches in high-skill participants thus accounted for their more stable arithmetic performance.

The next experiment was designed to test whether the effects observed in addition problems can be generalized to multiplication problems. Since problem-answer associations are generally less peaked for multiplication than for addition, we predict that people will not as easily change strategies for multiplication as for addition.

Experiment 2: Multiplication

Method

Participants. Forty-one subjects participated in the present experiment (22 men and 19 women; mean age 19 years 6 months). All subjects were first-year psychology students at Ghent University who participated for course requirements and credits.

The multiplication task. The multiplication problems consisted of two one-digit numbers (e.g., 6 x 7). Problems involving 0 or 1 as an operand (e.g., 5 x 0, 1 x 4) and tie problems (e.g., 3 x 3) were excluded. Since commuted pairs (e.g., 9 x 4 and 4 x 9) were considered as two different problems, this resulted in 56 multiplication problems (ranging from 2 x 3 to 9 x 8). If the product of both operands was smaller than 25, the problem was defined as small. If the product of both operands was larger than 25, the problem was defined as large.

Procedure. The procedure of Experiment 2 was completely identical to the procedure used in Experiment 1, with as only exception the explanation of the different strategy types: (1) Retrieval: You solve the problem by remembering or knowing the answer directly from memory. It means that you know the answer without any additional processing. For example: you know that 5 x 6 = 30 because 30 “pops into your head”. (2) Procedural: It means that you don’t know the answer without any additional processing. One possibility is that you solve the problem by counting a certain number of times to get the answer (i.e., reciting the tables of multiplication). For example: 4 x 7 = 7… 14… 21… 28 or 5 x 3 = 5… 10… 15. Another possibility is that you solve the problem by referring to related operations or by deriving the answer from known facts. You change the presented problem to take advantage of a known arithmetical fact. For example: 9 x 8 = (10 x 8) – 8 = 80 – 8 = 72 or 6 x 7 = (6 x 6) + 6 = 36 + 6 = 42. (3) Other: You solve the problem by a strategy unlisted here, or you do not know what strategy you used to solve the problem. For example: guessing.

The French Kit. The scores on the French Kit (median = 29) were again used to divide the group of participants into two subgroups. The low-skill participants’ score was below or equal to the median score (*N* = 22; mean = 24; range = 16 to 29), whereas the high-skill participants’ score was above the median score (*N* = 19; mean = 38, range = 30 to 53).

Results and discussion

Of all trials, 11% was spoiled due to failures of the sound-activated relay. Since all these invalid trials returned at the end of the block, most of them were recovered from data loss, which reduced the trials due to failures of the sound-activated relay to 2%. All trials on which participants reported having used a strategy ‘Other’ (1%) were deleted as well. All data were analyzed on the basis of the multivariate general linear model, and all reported results were significant at $p < .05$, unless mentioned otherwise. The results section is organized in the same four parts as in Experiment 1.

Secondary task performance. An ANOVA was run on the CRT task’s RTs (of correctly answered trials only); with Skill as between-subjects factor and repeated measures on Condition (single, naming, accuracy, accuracy/speed, speed; see Table 1). The significant effect of Condition, $F(4,34) = 20.01$, $MSe = 3241$, indicated that participants responded faster in the single task condition (577 ms) than in the condition in which they only had to read the answers off the screen (660 ms), $F(1,39) = 9.19$. Response times in the latter condition were faster than those observed under accuracy (686ms), accuracy/speed (682 ms), and speed instructions, (690 ms), $F(1,39) = 3.42$ ($p = .07$). Response times did not differ across the latter three dual-task conditions, $F<1$; and the effect of Skill did not reach significance, $F<1$. The Condition x Skill interaction was significant, $F(4,36) = 2.69$, $MSe = 3241$. In the single task condition, low-skill participants (604 ms) were slower than high-skill participants (550 ms), $F(1,39) = 4.70$, whereas there was no difference between high- and low-skill participants in the other conditions (each $p > .30$). Hence, high-skill participants have more working-memory resources available than low-skill participants, but this advantage moves to the primary task under dual-task conditions.

Primary task performance: speed. A 2 x 3 x 2 x 2 ANOVA was run on RTs (of correctly solved problems only, see Table 4) with Skill (low vs. high) as between-subjects factor and Instruction (accuracy vs. speed vs. accuracy/speed), Load (no-load vs. load), and Problem size (small vs. large) as within-subject factors. The main effect of Instruction, $F(2,38) = 9.74$, MSe = 373171, indicated that participants were slower under accuracy instructions (1456 ms) than under speed/accuracy instructions (1278 ms) $F(1,39) = 7.50$, and slower under accuracy/speed instructions than under speed instructions (1125 ms), $F(1,39) = 5.87$. Participants were also faster on small problems (955 ms) than on large problems (1619 ms), $F(1,39) = 83.16$, MSe = 648532, and faster in no-load conditions (1191 ms) than under working-memory load (1382 ms), $F(1,39) = 30.87$, MSe = 145167. High-skill participants (1043 ms) were faster than low-skill participants (1530 ms), $F(1,39) = 21.30$, MSe = 1358966. Skill further interacted with Load, $F(1,39) = 5.10$, MSe = 145167 and with Size, $F(1,39) = 11.48$, MSe = 648532. Load effects and problem size effects were higher for low-skill participants (269 ms and 911 ms) than for high-skill participants (113 ms and 417 ms). The other effects did not reach significance.

A similar 2 x 3 x 2 x 2 ANOVA was run on pure *retrieval* RTs (of correctly solved problems only; see Table 5). All main effects were significant, $F(2,38) = 11.43$, MSe = 120422 for Instruction, $F(1,39) = 113.31$, MSe = 173097 for Problem size, $F(1,39) = 35.97$, MSe = 131638 for Load, and $F(1,39) = 20.38$, MSe = 728327 for Skill. Skill also interacted with Load, $F(1,39) = 7.32$, MSe = 131638, and with Size, $F(1,39) = 11.31$, MSe = 173097. Hence, even for pure retrieval RTs, load effects and problem size effects are higher for low-skill participants (285 ms and 527 ms) than for high-skill participants (108 ms and 274 ms). The interaction between Skill and Instruction did not reach significance (F<1).

Primary task performance: errors. A similar 2 x 3 x 2 x 2 ANOVA was run on error percentages (see Table 4). The main effect of Instruction, $F(2,38) = 22.26$, MSe = 35, indicated that people made more errors in the speed condition (9.4%) than in the accuracy/speed condition (6.3%), $F(1,39) = 3.76$ ($p = .06$), and more errors in the accuracy/speed condition than in the accuracy condition (5.0%), $F(1,39) = 21.52$. Participants were also more erroneous on large problems (11.3%) than on small problems (2.5%), $F(1,39) = 143.53$, $MSe = 66$, and more erroneous in load conditions (7.6%) than in no-load conditions (6.2%), $F(1,39) = 13.41$, $MSe = 16$. Low-skill participants were more erroneous (8.1%) than high-skill participants (5.7%), $F(1,39) = 3.90$ ($p = .06$) , $MSe = 181$.

The Skill x Instruction interaction indicated that low-skill participants were more influenced by instructions than were high-skill participants. More specifically, the difference in error rates between accuracy and speed instructions, was much larger in low-skill participants (6.2%) than in high-skill participants (2.4%), $F(1,39) = 8.15$. The Skill x Size interaction indicated that the problem size effect was larger in low-skill participants (10.6%) than in high-skill participants (7.0%), $F(1,39) = 6.24$, $MSe = 66$.

The same 2 x 3 x 2 x 2 ANOVA was run on error percentages of *retrieval* trials only (see Table 5). The main effects of Instruction, Load, and Size were significant, $F(2,38) = 25.34$, $MSe = 37$, $F(1,39) = 7.05$, $MSe = 18$, and $F(1,39) = 94.88$, $MSe = 65$. The main effect of Skill failed to reach significance, $F(1,39) = 3.05$, $MSe = 110$ ($p = .09$), but the Skill x Instruction interaction was significant, $F(2,38) = 6.65$, $MSe = 37$. Low-skill participants were more influenced by instructions than high-skill participants. Across the different instructions, retrieval error rates ranged from 2.9% to 9.7% for low-skill participants but only from 3.9% to 5.8% for high-skill participants. The Skill x Size and Skill x Load interactions did not reach significance though (both Fs<1).

Primary task performance: strategy choices. The same 3 x 2 x 2 ANOVA was run on percentages retrieval use (correctly solved problems only; see Table 4). The main effect of

Instruction, $F(2,38) = 7.01$, $MSe = 165$, indicated that the retrieval strategy was chosen more frequently under speed and accuracy/speed instructions (85.9% and 83.9%) than under accuracy instructions (81.3%), $F(1,39) = 12.12$ and $F(1,39) = 4.26$, respectively. Further, participants chose the retrieval strategy more frequently on small problems (93.3%) than on large problems (74.1%), $F(1,39) = 62.79$, $MSe = 715$ and more frequently in load conditions (85.0%) than in no-load conditions (82.4%), $F(1,39) = 10.38$, $MSe = 84$. The main effect of Skill was not significant, $F<1$, and Skill did not interact with other variables (each $p > .05$), which indicated that the differences in multiplication efficiency between high-skill and low-skill participants cannot be due to different strategy choices between both types of participants. As was already clear from the analyses on retrieval efficiency, high-skill participants are more efficient than low-skill participants because they are more flexible in adapting their retrieval efficiency.

Discussion. Instruction had an effect on participants' accuracies, RTs, and strategy choices. As in Experiment 1, participants were more accurate under accuracy instructions and faster under speed instructions. The tradeoff between speed and accuracy was also present. The changes in accuracy and speed were accompanied by strategy selection effects. People switched to the faster retrieval strategy under speed conditions, but reverted to slower non-retrieval strategies under accuracy instructions.

The main effect of skill indicated that high-skill participants were faster and more accurate than low-skill participants. The accuracies of high-skill participants were also more stable than those of low-skill participants. That is, the error increase under speed instructions was much larger for low-skill than for high-skill participants (see also Smith-Chant & LeFevre, 2003). Interestingly, we also observed larger load effects and larger problem size effects in low-skill than in high-skill participants. In contrast to Experiment 1, these differential effects between high- and low skill participants were not caused by different strategy choices. They were rather related to different strategy efficiencies between both participant groups. As already outlined in the introduction, people have less strong problem-answer associations for multiplication, which implies less highly peaked distributions of associations and thus less room for change. Though, the "peakedness" is higher for high-skill than for low-skill participants, which explains the lower instruction, load, and size effects in high-skill than in low-skill participants.

Insert Tables 4 and 5 about here

General Discussion

In the present study, we tested the effects of instruction and working-memory load on high-skill and low-skill participants when solving simple addition problems (Experiment 1) and simple multiplication problems (Experiment 2). Instruction effects were observed for RTs, accuracies, and strategy choices. When instructed to respond fast, participants were fast indeed but they abandoned accuracy. Similarly, when instructed to respond accurately, participants were accurate indeed but they abandoned speed. When instructed to respond fast *and* accurately, participants' performance was in-between. Participants were thus flexible, but they were not capable to answer accurately *and* fast – there was always a tradeoff between accuracy and speed. In both experiments, participants' strategy choices were also affected by instruction effects. Retrieval was used more frequently under speed instructions than under accuracy instructions.

In the ASCM (Siegler & Jenkins, 1989: Siegler & Shipley, 1995), the increased retrieval use under speed instructions can be explained by changing confidence criteria. A confidence criterion determines how sure one must be before stating a retrieved answer. As soon as an activated number node in long-term memory exceeds this criterion, the answer is retrieved. If none of the activated number nodes exceeds this criterion, a non-retrieval strategy is used to solve the problem. Hence, when participants under speed instructions lowered their confidence criterion, they had to be less sure about the answer before choosing the retrieval strategy. As a consequence, the frequency of retrieval increased, resulting in faster but more erroneous answers. Also note that a change in the search length criterion would not lead to the same conclusions. When people would decide to lower their search length criterion under speed instructions, they would revert in more frequent non-retrieval use and hence slower responding. The rather contra-intuitive decision to higher the search length criterion under speed instructions would indeed lead to more frequent retrieval use, but the retrieval efficiency itself would be lower. This was not the case, since the retrieval-only analyses showed faster retrieval use under speed instructions.

Instruction effects differed across high- and low-skill participants. In the addition experiment, high-skill participants reacted adaptively to the instructions by changing their strategy choices. Because distributions of associations for addition are generally more highly peaked than for multiplication, switching strategies was easier in Experiment 1 as compared to Experiment 2. The higher start position of the high-skill participants' confidence criterion also tolerated a greater range to change the criterion adaptively according to the situation. Hence, the combination of highly peaked distributions of associations and high confidence criteria allowed high-skill participants to safely lower their confidence criterion without loosing too much speed or accuracy. Because lowering the confidence criterion results in more frequent retrieval use, high-skill participants applied this technique under load conditions and for large problems – which was adaptive because the retrieval strategy is less resource-demanding than other, non-retrieval strategies. These adaptive strategy switches also explain why the load effect (on RTs) was smaller for high-skill than for low-skill participants. Low-skill participants have less room to be adaptive because their distributions of associations are less highly peaked and their confidence criteria are lower. Skill and adaptivity are thus closely interconnected: the larger the amount of arithmetic knowledge, the more room for adaptive behavior. Because low-skill participants do not know the answers to certain arithmetic problems, they cannot change their strategy choices and – as a consequence – they are labeled as less adaptive.

In the multiplication experiment, high-skill participants did not react adaptively to the instructions by showing more frequent retrieval use. Because distributions of associations are less highly peaked and confidence criteria are less high for multiplication than for addition, there was less room to change strategy choices. However, high-skill participants did change their strategy efficiencies: the execution of the retrieval strategy itself became more efficient. In terms of the ASCM, this means that – for multiplication problems – high-skill participants have stronger problem-answer associations and higher confidence criteria. That is, even though the high-skill participants' confidence criterion was stricter, the strong problem-answer associations enabled them to use the fast retrieval strategy as frequently as low-skill participants, who had a less strict confidence criterion but also much weaker problem-answer associations. Hence, despite equally frequent retrieval use, the execution of the retrieval strategy itself was more efficient (i.e., faster and less erroneous) for high-skill than for than low-skill participants. Low-skill participants might not have been able to respond to the instructions in a similar way because they were already performing as efficiently as they could.

Both types of adaptive behavior (changes in strategy choices in Experiment 1 and changes in strategy efficiency in Experiment 2) resulted in a more stable pattern in terms of RTs and accuracies. Low-skill participants, in contrast, did not change their behavior as adaptively as the high-skill participants, resulting in a less stable pattern in terms of RTs and accuracies. More specifically, they abandoned speed under accuracy instructions (Experiment 1) and abandoned accuracy under speed instructions (Experiment 2). Similar results were obtained by Smith-Chant and LeFevre (2003). However, as noted in the introduction, Smith-Chant and LeFevre did not test the role of working memory. And working memory does matter, as discussed below.

In both experiments, we observed significant load effects on RTs and accuracies. Significant amounts of working-memory resources are thus required in order to solve simple addition and multiplication problems with a reasonable level of speed and accuracy (see DeStefano & LeFevre, 2004, for review). High-skill participants also needed fewer working memory resources than did low-skill participants, probably because they already had more working-memory resources available to start with (cf. secondary task results Experiment 2). Furthermore, high-skill participants are also frequent users of the retrieval strategy, which is less resource-demanding than other, non-retrieval strategies (e.g., Hecht, 2002; Imbo & Vandierendonck, 2007a,b,c; Seyler, Kirk, & Ashcraft, 2003).

In Experiment 1, load effects were not only observed on RTs and accuracies, but also on percentages retrieval use. This was not only unexpected; it is also in disagreement with previous studies (Hecht 2002; Imbo & Vandierendonck, 2007a,b,c; but see Imbo, Duverne, & Lemaire, 2007). High-skill participants under load conditions (i.e., with fewer working-memory resources left) switched to the less-demanding retrieval strategy rather than continue using the more-demanding non-retrieval strategies. Such strategy switches as a result of a decrease in the available working-memory resources cannot be accounted for by the ASCM. In the ASCM, strategy selection is based on simple basic processes such as activation weighting and association strengthening and not on conscious, deliberate, or metacognitive processes requiring working memory resources. However, the more recent strategy choice and discovery simulation model (SCADS; Shrager & Siegler, 1998; Siegler & Arraya, 2005) includes a metacognitive system with an attentional spotlight. This attentional spotlight allocates resources to the execution of poorly learned strategies. When enough attentional resources are available, they are used to discover new strategies or to interrupt the execution of an ongoing strategy. The attentional spotlight thus plays a role in strategy execution and strategy discovery, but not in strategy selection. We suggest that future adaptations of the SCADS should allow the attentional spotlight to interfere with the strategy selection process as well (e.g., under load conditions).

Our results cannot only be framed within the ASCM, which is specific for the mathematical domain. At a more general level, our data are also in line with other models of strategy selection, such as the Adaptive Decision Maker (Payne et al., 1993). Payne and colleagues assume that people's strategy choices are affected by three factors: characteristics of the different strategies, individual differences, and contextual factors. Our study clearly showed the influence of all three factors. First, the characteristics of retrieval versus non-retrieval strategies (i.e., their efficiency and their reliance on working-memory resources) urged participants to use the former strategy more frequently than the latter one. Second, strategy choices also differed as a function of the only individual difference under investigation, mathematical skill: retrieval was used more frequently by high-skill than by low-skill participants. And finally, instructional demands such as the requirement to answer quickly rather than accurately also affected participants' strategy choices. Interestingly, these three factors also interacted in the process of adaptive strategy selection. Most interestingly, high-skill participants responded more adaptively to changing situations than did low-skill participants. This is in agreement with Payne and colleagues as well, since these authors argue that people adapt their strategy choices to contextual factors because the relative advantages and disadvantages of each strategy change according to the situation. Indeed, we observed that high-skill participants switched to the less resource-demanding retrieval strategy under load conditions. Stated differently, one of the advantages of the retrieval strategy (i.e., its very low reliance on working-memory resources) is relatively more important under load conditions than under no-load conditions.

The observation of adaptivity differences between high- and low-skill participants is also in agreement with a study by Schunn and Reder (2001), who claim that the difference in strategy adaptivity has for long been ignored as an important factor in individual differences. However, Schunn and Reder specifically focused on adaptivity to changing success base rates (see also Reder & Schunn, 1999; Schunn & Reder, 1998), whereas we rather focused on adaptivity to changing instructions and changing load conditions. Further research is thus needed to investigate (a) which variables people take into account when switching strategies (e.g., success base rates, the amount of working memory load, the instructions imposed by the experimenter, the amount of stress, the complexity of the task, et cetera), (b) why some people are more adaptive than others. Although the results obtained in the present study suggest that adaptivity is related to arithmetic skill, it is way too early to infer a causal relationship. It is not clear whether a more adaptive behavior causes higher levels of arithmetic skill or whether it is rather the other way around. Anyhow, there are many questions about adaptivity that still have to be solved – and the conclusions are not only relevant for theoretical models, but also for everyday life. In the field of organizational psychology for example, it might be advantageous to hire adaptive individuals for jobs in which the task environment changes frequently and rapidly. More general, because performance is so highly related to strategy adaptivity (with higher and/or more stable scores for adaptive than for non-adaptive persons), it can be concluded that strategy adaptivity is a very important factor – at both the theoretical and applied level of cognitive processes.

To conclude, when asked to respond as fast and/or as accurately as possible, people are quite well in obeying these instructions. However, high-skill participants were better in adaptively adjusting their behavior, resulting in a highly efficient and stable pattern of arithmetic performance (RTs and accuracies). Low-skill participants, in contrast, were way more submissive and did not change their strategy choices when the amount of available working-memory resources was reduced. This resulted in a less stable pattern of arithmetic performance: low-skill participants abandoned speed for accuracy and accuracy for speed – and these effects were even more harmful for large problems. Stated differently, high-skill participants *cope* with changing situations (e.g., a higher working memory load or other instruction conditions), whereas low-skill participants *undergo* changing situations. The ASCM can account for both patterns of behavior, that is, the flexible, adaptive behavior of high-skill participants and the less flexible, non-adaptive behavior of low-skill participants.

Footnotes

1. For Experiment 1 (Addition), independent sample *t*-tests showed that high-skill and low-skill participants did not differ in male/female ratio, math anxiety, or math experience, each $t(38)<1.0$. High-skill and low-skill participants did differ in calculator use, however, $t(38) = 2.25$. Low-skill participants reported more frequent calculator use (3.3) than did high-skill participants (2.5). For Experiment 2 (Multiplication), independent sample *t*-tests showed that high-skill and low-skill participants did not differ in male/female ratio, math anxiety, or calculator use, each $t(39)<1.8$. High-skill and low-skill participants did differ in math experience, however, $t(39)= 2.80$. The amount of math experience was higher in high-skill participants (5.16) than in low-skill (3.86) participants.

2. We acknowledge that the only method to achieve 100% unbiased strategy efficiency data is the choice/no-choice method (Siegler & Lemaire, 1997). In this method, there is not only a choice condition, in which participants are allowed to choose among several solution strategies. This method includes several no-choice conditions as well, in which participants are asked to use one single strategy to solve all problems. The accuracy and speed data obtained in these no-choice conditions then provide unbiased strategy efficiency data. The reason why the choice/no-choice method was not applied in the present experiment is its incompatibility with accuracy and speed instructions. It makes little sense to ask participants to respond “as fast as possible” while requesting them to use – for example – a counting strategy. Such incompatible requirements may cause discrepancies between the used strategy (e.g., retrieval), on the one hand, and the reported strategy (e.g., counting), on the other. These inconsistencies may then, in turn, bias the results (see also Kirk & Ashcraft, 2001).

References

Ashcraft, M.H. (1995). Cognitive psychology and simple arithmetic: A review and summary of new directions. *Mathematical Cognition, 1*, 3-34.

Baddeley, A.D., & Logie, R.H. (1999). Working memory: The multi-component model. In A. Miyake & P. Shah (Ed.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28-61). New York: Cambridge University Press.

DeStefano, D., & LeFevre, J.-A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*, 353-386.

Engle, R.W., Tuholski, S.W., Laughlin, J.E., & Conway, A.R. (1999). Working memory, STM, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General, 128*, 309-331.

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

Hecht, S.A. (2002). Counting on working memory in simple arithmetic when counting is used for problem solving. *Memory & Cognition, 30*, 447-455.

Imbo, I., Duverne, S., & Lemaire, P. (2007). Working memory, strategy execution, and strategy selection in mental arithmetic. *Quarterly Journal of Experimental Psychology, 60*, 1246-1264.

Imbo, I., & Vandierendonck, A. (2007a). The role of phonological and executive working-memory resources in simple-arithmetic strategies. *European Journal of Cognitive Psychology, 19*, 910-933.

Imbo, I., & Vandierendonck, A. (2007b). Do multiplication and division strategies rely on executive and phonological working-memory resources? *Memory & Cognition, 35*, 1759-1771.

Imbo, I., & Vandierendonck, A. (2007c). Strategy use in elementary-school children: Working memory and individual differences. *Journal of Experimental Child Psychology*, *96*, 284-309.

Imbo, I., & Vandierendonck, A. (2008). Effects of problem size, operation, and working-memory span on simple-arithmetic strategies: Differences between children and adults? *Psychological Research, 72*, 331-346.

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, & Cognition, 27*, 157-175.

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., 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.

Miyake, A. (2001). Individual differences in working memory: Introduction to the special section. *Journal of Experimental Psychology: General*, *130*, 163-168.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge: University Press.

Reder, L. M., & Schunn, C. D. (1999). Bringing together the psychometric and strategy worlds: Predicting adaptivity in a dynamic task. In D. Gopher & A. Koriat (Eds.), *Cognitive* *regulation of performance: Interaction of theory and application. Attention and Performance XVII* (pp. 315-342). Cambridge, MA: MIT Press.

Schunn, C. D., & Reder, L. M. (1998). Individual differences in strategy adaptivity. In D. L. Medin (Ed.), *The psychology of learning and motivation* (pp. 115-154). York, PA: Academic Press.

Schunn, C. D., & Reder, L. M. (2001). Another source of individual differences: Strategy adaptivity to changing rates of success. *Journal of Experimental Psychology: General, 130*, 59–76.

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., & Araya, R. (2005). A computational model of conscious and unconscious strategy discovery. *Advances in Child Development and Behavior, 33*, 1-42.

Siegler, R.S., & Jenkins, E.A. (1989). *How children discover new strategies*. Hillsdale, NJ: Erlbaum.

Siegler, R.S., & Lemaire, P. (1997). Older and younger adults' strategy choices in multiplication: Testing predictions of ASCM using the choice/no-choice method. *Journal of Experimental Psychology: General, 126*, 71-92.

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

Smith-Chant, B.L., & LeFevre, J.-A. (2003). Doing als they are told and telling like it is: Self-reports in mental arithmetic. *Memory & Cognition, 31*, 516-528.

Szmalec, A., Vandierendonck, A., & Kemps, E. (2005). Speeded Decision-Making as a Component of Executive Functioning. *Memory & Cognition*, *33*, 531-541.

Table 1

Mean response times (ms) on the CRT task as a function of Arithmetic skill and Condition. Standard errors between brackets.

	Experiment 1 (Addition)		Experiment 2 (Multiplication)	
	Low skill	High skill	Low skill	High skill
Single	582 (19)	568 (20)	604 (19)	550 (20)
Naming	677 (13)	682 (13)	677 (16)	643 (17)
Accuracy	715 (16)	706 (16)	670 (16)	701 (15)
Accuracy/Speed	691 (15)	689 (16)	679 (14)	694 (17)
Speed	671 (15)	724 (16)	676 (17)	696 (19)

Table 2

Mean RTs (ms), error rates (%), and retrieval usage (%) for addition problems (Experiment 1) as a function of Arithmetic skill, Problem size, Instruction, and Load. Standard errors between brackets.

	RT (ms)	Accuracy		Accuracy/Speed		Speed	
		No load	Load	No load	Load	No load	Load
Low-skill	Small	848 (40)	1160 (88)	716 (36)	959 (57)	695 (31)	839 (39)
	Large	1133 (55)	1522 (89)	1018 (51)	1212 (62)	922 (42)	1084 (51)
High-skill	Small	723 (42)	871 (93)	652 (38)	838 (60)	646 (33)	797 (41)
	Large	1000 (58)	1118 (45)	847 (54)	1022 (66)	834 (45)	990 (53)
	Error rate (%)	Accuracy		Accuracy/Speed		Speed	
		No load	Load	No load	Load	No load	Load
Low-skill	Small	0.9 (0.4)	1.9 (0.7)	1.6 (0.6)	2.1 (0.7)	3.0 (0.8)	4.3 (0.9)
	Large	3.4 (1.0)	3.5 (0.8)	4.5 (1.3)	6.8 (1.6)	9.3 (1.6)	9.2 (1.5)
High-skill	Small	0.6 (0.5)	0.6 (0.7)	0.3 (0.6)	0.9 (0.7)	1.9 (0.9)	0.9 (0.9)
	Large	1.8 (1.0)	2.7 (0.8)	3.7 (1.3)	3.5 (1.7)	3.9 (1.7)	6.0 (1.6)
	Retrieval use (%)	Accuracy		Accuracy/Speed		Speed	
		No load	Load	No load	Load	No load	Load
Low-skill	Small	91.0 (3.2)	89.7 (3.4)	95.3 (3.2)	89.9 (3.4)	93.7 (2.7)	89.6 (3.3)
	Large	58.5 (4.9)	62.2 (5.6)	63.0 (5.0)	61.7 (5.1)	65.9 (5.3)	67.8 (5.2)
High-skill	Small	98.1 (3.4)	94.3 (3.6)	96.8 (3.4)	93.1 (3.5)	96.3 (2.8)	94.9 (3.5)
	Large	56.4 (5.1)	65.6 (5.9)	61.0 (5.3)	65.8 (5.3)	58.8 (5.6)	70.3 (5.4)

Table 3

Mean *retrieval* RTs (ms) and *retrieval* error rates (%) for addition problems (Experiment 1) as a function of Arithmetic skill, Problem size, Instruction, and Load. Standard errors between brackets.

	RT (ms)	Accuracy		Accuracy/Speed		Speed	
		No load	Load	No load	Load	No load	Load
Low-skill	Small	698 (35)	930 (58)	822 (37)	1119 (88)	676 (29)	801 (37)
	Large	869 (37)	1063 (56)	922 (36)	1308 (81)	805 (33)	974 (38)
High-skill	Small	650 (37)	840 (61)	723 (39)	859 (93)	637 (31)	788 (39)
	Large	733 (39)	921 (59)	850 (38)	983 (85)	746 (35)	892 (40)

	Error rate (%)	Accuracy		Accuracy/Speed		Speed	
		No load	Load	No load	Load	No load	Load
Low-skill	Small	0.6 (0.3)	0.8 (0.3)	0.5 (0.3)	1.2 (0.5)	2.2 (0.8)	3.0 (0.8)
	Large	2.9 (1.1)	1.3 (0.7)	1.3 (0.7)	0.7 (0.5)	8.2 (1.5)	6.6 (1.5)
High-skill	Small	0.3 (0.4)	0.3 (0.4)	0.3 (0.4)	0.3 (0.6)	1.3 (0.8)	0.6 (0.8)
	Large	2.8 (1.2)	2.6 (1.6)	0.8 (0.7)	1.5 (0.5)	3.5 (1.6)	5.2 (1.5)

Table 4

Mean RTs (ms), error rates (%), and retrieval usage (%) for multiplication problems (Experiment 2) as a function of Arithmetic skill, Problem size, Instruction, and Load. Standard errors between brackets.

	RT (ms)	Accuracy		Accuracy/Speed		Speed	
		No load	Load	No load	Load	No load	Load
Low-skill	Small	1063 (41)	1291 (62)	946 (38)	1165 (69)	877 (31)	1105 (58)
	Large	2184 (174)	2417 (178)	1845 (168)	2182 (175)	1456 (97)	1827 (115)
High-skill	Small	844 (43)	951 (67)	794 (41)	879 (74)	720 (33)	820 (63)
	Large	1337 (187)	1564 (191)	1177 (181)	1238 (189)	1047 (104)	1149 (123)
	Error rate (%)	Accuracy		Accuracy/Speed		Speed	
		No load	Load	No load	Load	No load	Load
Low-skill	Small	1.3 (0.7)	2.5 (0.9)	1.6 (0.5)	3.7 (0.8)	3.1 (0.7)	4.5 (0.9)
	Large	8.2 (1.4)	9.2 (1.5)	12.4 (1.8)	12.1 (1.6)	17.2 (1.9)	21.3 (2.1)
High-skill	Small	2.2 (0.7)	2.7 (1.0)	1.2 (0.5)	2.3 (0.9)	3.0 (0.7)	2.0 (1.0)
	Large	5.4 (1.5)	8.7 (1.6)	8.9 (2.0)	8.0 (1.7)	10.3 (2.1)	13.7 (2.3)
	Retrieval use (%)	Accuracy		Accuracy/Speed		Speed	
		No load	Load	No load	Load	No load	Load
Low-skill	Small	89.4 (2.8)	92.5 (3.2)	91.4 (2.6)	94.5 (2.1)	92.0 (2.2)	93.0 (2.9)
	Large	66.1 (5.3)	77.0 (5.2)	68.9 (5.8)	74.6 (0.5)	77.4 (5.0)	78.3 (5.0)
High-skill	Small	95.4 (3.0)	93.2 (3.5)	94.0 (2.8)	94.5 (2.3)	95.6 (2.4)	93.8 (3.1)
	Large	67.9 (5.7)	69.1 (5.6)	72.9 (6.3)	80.3 (5.4)	77.4 (5.4)	79.6 (5.3)

Table 5

Mean *retrieval* RTs (ms) and *retrieval* error rates (%) for multiplication problems (Experiment 2) as a function of Arithmetic skill, Problem size, Instruction, and Load. Standard errors between brackets.

	RT (ms)	Accuracy		Accuracy/Speed		Speed	
		No load	Load	No load	Load	No load	Load
Low-skill	Small	983 (39)	1243 (60)	893 (34)	1126 (61)	836 (29)	1078 (54)
	Large	1566 (96)	1924 (122)	1362 (80)	1623 (103)	1243 (70)	1601 (87)
High-skill	Small	806 (42)	923 (65)	751 (37)	842 (65)	705 (32)	799 (58)
	Large	1113 (103)	1292 (131)	1002 (86)	1085 (111)	947 (75)	1031 (94)
	Error rate (%)	Accuracy		Accuracy/Speed		Speed	
		No load	Load	No load	Load	No load	Load
Low-skill	Small	0.9 (0.5)	1.7 (0.6)	0.8 (0.4)	3.3 (0.8)	2.4 (0.6)	3.8 (0.8)
	Large	5.0 (1.3)	3.9 (1.7)	9.5 (2.1)	9.0 (1.3)	14.0 (1.9)	18.1 (2.1)
High-skill	Small	0.9 (0.5)	1.6 (0.6)	0.9 (0.4)	1.3 (0.9)	1.7 (0.6)	1.0 (0.9)
	Large	2.8 (1.4)	9.1 (1.8)	8.2 (2.3)	5.1 (1.4)	9.4 (2.0)	11.0 (2.2)

Figure 1

Response times (ms) for Experiment 1 (Addition) as a function of Skill, Instruction, and Working-memory load.

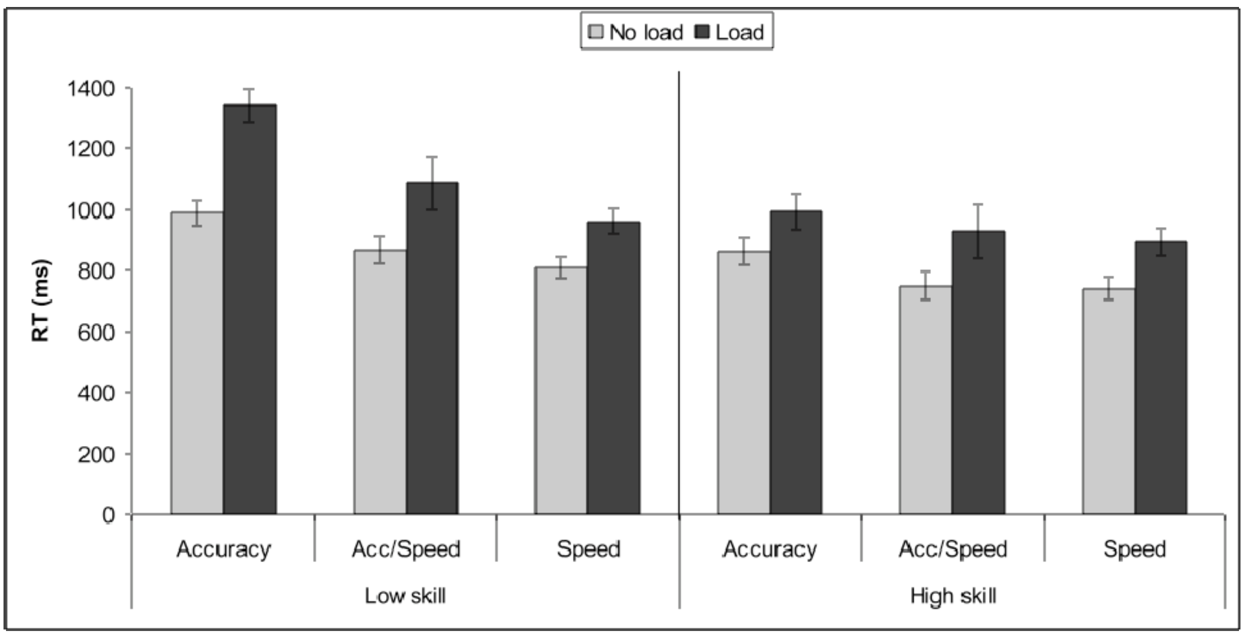

Figure 2

Retrieval use (%) for Experiment 1 (Addition) as a function of Skill, Problem size, and Working-memory load.