Memory & Cognition
2001, 29 (2), 267-273

Verifying simple arithmetic sums and products: Are the phonological loop and the central executive involved?

STIJN DE RAMMELAERE, ELS STUYVEN, and ANDRÉ VANDIERENDONCK
Ghent University, Ghent, Belgium

In two experiments, we investigated the role of the phonological loop and the central executive in the verification of the complete set of one-digit addition (Experiment 1) and multiplication (Experiment 2) problems. The focus of the present study was on the contradictory results concerning the contribution of the phonological loop in the verification of true problems (e.g., 8 + 4 = 12 or 4 × 6 = 24) reported until now. The results revealed that this slave system is not involved in verifying simple arithmetic problems, in contrast to the central executive. Furthermore, our results indicated that the split effect is due to the use of two different arithmetic strategies.

The purpose of the present study is to further explore the relation between mental arithmetic and working memory. Hitch (1978) is commonly referred to as one of the first researchers to investigate the contribution of working memory to mental arithmetic. Hitch presented such problems as 425 + 63 and found that the most important sources of errors are (1) forgetting partial results of calculation (such as 5 + 3 in our example) and (2) forgetting initial information. Thus, Hitch suggested that working memory made a crucial contribution to mental arithmetic but did not investigate in detail which components of the multicomponent working memory model (Baddeley & Hitch, 1974) are involved. This model has already proven to be a useful conceptualization of working memory. According to its authors, working memory refers to temporary storage and processing in a variety of cognitive tasks and consists of three components. Two are specialized slave systems that are responsible for the processing of verbal (phonological loop) and visual/spatial (visuospatial sketch pad) information. The third component is the central executive, a modality-free limited-capacity attentional system that, among other things, monitors the allocation of mental resources to the two slave systems during cognitive tasks.

Logie, Gilhooly, and Wynn (1994) used this model to investigate the different components of working memory in complex arithmetic problems. They used a calculation task in which participants had to add a series of two-digit numbers while the different components of working memory were loaded. Logie et al. concluded that (1) the role of the phonological loop is probably to keep track of running totals and to maintain accuracy in calculation and (2) the role of the central executive is most likely to perform the calculations required for mental addition and to produce approximately correct answers. Because Logie et al. presented a series of numbers and the participants had to state their responses only at the end of the series, the question of whether the central executive and the phonological loop are also involved in single sums (e.g., 18 + 17 = ? or 3 + 5 = ?) remained unsolved. The study of Ashcraft, Donley, Halas, and Vakali (1992) was the first to suggest that even for relatively simple arithmetic, access to arithmetic facts and their manipulation are not entirely automatic and probably rely on working memory resources.

Lemaire, Abdi, and Fayol (1996) investigated the role of working memory in single, simple arithmetic problems. These researchers used the verification task (e.g., 8 + 4 = 12 True? False?) and investigated only one-digit numbers from 2 to 9. The false answers were created in two ways. First, the false answer in a *confusion problem* was the product (for addition) or the sum (for multiplication) of the two terms (e.g., 5 + 3 = 15 or 4 × 6 = 10). The false answer in a *nonconfusion problem*, on the contrary, was the product plus or minus 1 or the sum plus or minus 1 (e.g., 5 + 3 = 14 or 16; 4 × 6 = 9 or 11), in order to equate the splits. Lemaire et al. investigated a control condition, a condition with articulatory suppression, and a condition with random letter generation. The results revealed that a load on the central executive (by means of random letter generation) disrupted the verification of both true and false equations, whereas articulatory suppression had a detrimental effect only for the true arithmetic problems.

S.D.R. is a research assistant at the Fund for Scientific Research (Flanders, Belgium). The research reported in this article was supported by Ph.D. Grant BOF 011D0896 from the Research Council of Ghent University to E.S. and by Grant G001297 from the Fund for Scientific Research–Flanders to A.V. We are grateful to Marc Brysbaert, Wim Fias, Patrick Lemaire, and two anonymous reviewers for helpful comments. Correspondence concerning this article should be directed to S. De Rammelaere, Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium (e-mail: stijn.derammelaere@rug.ac.be).

In other words, this suggests that the central executive is crucial for both true and false equations, whereas the phonological loop is involved only in true problems.

De Rammelaere, Stuyven, and Vandierendonck (1999) replicated the study of Lemaire et al. (1996) but restricted the stimuli and investigated only addition. Some modifications were introduced. First, the stimuli were created in another way. Lemaire et al. used only (1) the sum or the sum plus or minus 1 (for multiplication) or (2) the product or the product plus or minus 1 (for addition) in order to get false answers. In this way, the false answers in their study were predominantly *extreme* ones (e.g., 7 + 8 = 55 or 56 or 57; 9 × 7 = 15 or 16 or 17), which are relatively easy to solve (*split* effect). In order to avoid a focus on extreme splits, De Rammelaere et al. combined the sums with the smallest split possible (+1) and with a larger but not extreme split (+5). Second, the random time interval generation (RIG; Vandierendonck, De Vooght, & Van der Goten, 1998) task was used as a new secondary task. In this task, participants are requested to tap a randomly spaced sequence of time intervals on a key so as to produce an unpredictable, random "rhythm." The requirement to be random and to avoid automaticity loads the central executive, whereas there is neither an empirical nor a logical ground to assume that there is interference with one of the known slave systems (in contrast to other secondary tasks, such as random letter generation, that also interfere with the phonological loop). This feature of the RIG task enables researchers to get more convincing evidence concerning a possible role of the central executive.

De Rammelaere et al. (1999) found that random letter generation and RIG interfered with the verification of both true and false problems, which replicated the crucial contribution of the central executive, as reported in the study of Lemaire et al. (1996). Another replicated finding was the outcome that articulatory suppression had no effect on false sums. However, the results concerning the role of the phonological loop in the verification of true problems were different: No effect of articulatory suppression on the true sums was found. This contradictory finding is important because it can shed light on the actual and very vivid debate as to whether basic arithmetic facts are stored in a language-dependent verbal form or not (e.g., Brysbaert, Fias, & Noël, 1998; Campbell, 1998; Dehaene, 1992; Dehaene & Cohen, 1995; Noël, Fias, & Brysbaert, 1997; Noël, Robert, & Brysbaert, 1998). The triple-code model of Dehaene (Dehaene, 1992; Dehaene & Cohen, 1995), for instance, assumes that arithmetic facts are stored in a verbal word frame, implying that a fact such as 2 × 3 = 6 cannot be retrieved unless the problem is coded into a verbal code "two times three. . . ." Interestingly, this model predicts that articulatory suppression, which makes verbal coding (almost) impossible, will interfere with true problems, but not with false ones, since it can be assumed that only true problems are stored. Although not interpreted as such, this was exactly what Lemaire et al. found. Since this finding was not replicated in the study of De Rammelaere et al., it is important to clarify the effect of articulatory suppression on the verification of basic arithmetic facts.

THE PRESENT STUDY

Before we will be able to reach more definitive conclusions, it is necessary to overcome some shortcomings of the study of De Rammelaere et al. (1999). First, only a subset of all possible one-digit sums was studied—namely, only 24 out of 64 possible single sums of the form $a + b$, where a and b are numbers from 2 to 9. Moreover, for the false answers, only positive splits (+1 and +5) were used.

The main goal of the present study was to further explore the contradictory findings about the effect of articulatory suppression. Furthermore, we wanted to test whether the conclusions of De Rammelaere et al. (1999), which were based on a limited stimulus set, can be generalized. In Experiment 1, participants verified the whole set of one-digit sums in a control condition, a condition with articulatory suppression, and an RIG condition.

EXPERIMENT 1

Method

Participants. Thirty first-year (mean age, 18.9 years) psychology students (4 males, 26 females) at Ghent University participated as a course requirement.

Stimuli. The stimuli were single sums presented in standard form (i.e., $a + b = c$); symbols and numbers were separated by spaces equal to the width of one character. The terms a and b were always one-digit numbers from 2 to 9. The sums were combined with (1) the correct answer (e.g., 8 + 4 = 12), (2) a split of ±1 (e.g., 8 + 4 = 11 or 13), and (3) a split of ±9 (e.g., 8 + 4 = 3 or 21). Thus, the split of +5 of De Rammelaere et al. (1999) was replaced by a split of ±9 in order to investigate additional false splits (i.e., between the +5 of De Rammelaere et al. and the extreme splits of Lemaire et al., 1996). Note that both positive and negative splits were studied.

When the numbers from 2 to 9 are investigated, a total of 64 combinations of the form $a + b$ is possible. Every combination was studied and was combined once with a small split (±1), once with a large split (±9), and twice with the correct solution so as to equate the numbers of true and false answers. As a result 256 (64 × 4) combinations of the form $a + b = c$ were created. These 256 combinations were randomly divided into three blocks (and completed with some deliberate filler items), so that the proportions of (1) true versus false sums, (2) small versus large splits, and (3) positive versus negative splits were equivalent in each block.

A split of +1 was assigned to a random half of the sums with a split of ±1; a split of −1 was allocated to the other half. An analogous procedure was followed with the split of ±9, but the sums with a correct answer of 9 or smaller always got a positive split, so the c term was never smaller than 1. The c term was never the product of $a \times b$, so the associative confusion effect could not play a role in the verification of the sums.

Each of the three blocks consisted of 88 trials (44 correct, 22 small split, 22 large split), with the filler items included. Within each block, the sequence of the sums was randomized, with the two following constraints: (1) A maximum of four successive trials could require the same response, and (2) if, for instance, 9 + 4 = 13 was presented, it was not possible that the following trial was 9 + 4 nor 4 + 9, combined with whatever answer.

Table 1
Experiment 1: Mean Latencies (RTs, in Milliseconds), Standard Deviations, and Accuracy for the Combinations of Load, Size, and Split

		CON			AS			RIG		
Size	Split	RT	*SD*	AC	RT	*SD*	AC	RT	*SD*	AC
Small	true	957	228	.94	971	234	.91	1,117	252	.92
	small	1,115	295	.94	1,116	254	.95	1,282	366	.88
	large	959	195	.98	1,016	246	.96	1,193	364	.98
Large	true	1,117	254	.92	1,148	294	.90	1,295	342	.90
	small	1,353	315	.85	1,340	359	.84	1,448	478	.89
	large	1,036	247	.98	1,053	245	.97	1,193	250	.96

Note—CON, control; AS, articulatory suppression; RIG, random time interval generation; AC, accuracy.

Procedure and Design. The stimuli were presented horizontally in the center of a computer screen, in yellow on a black background. The equations remained on until the participant responded or until 10 sec elapsed without response. The participants were instructed to solve the sums as accurately and as fast as possible by pressing the appropriate key. The left and the right buttons of the mouse were designated as true or false. For half of the participants the left button was designated as true, and for the other half it was designated as false. All the participants were instructed to use their index finger and the middle finger of their right hand to press these keys. The intertrial interval was 1 sec. Each subject participated in every condition. There were three conditions: control (CON), articulatory suppression (AS), and RIG.

In the CON condition, the participants solved the sums without a secondary task. The AS condition required the participants to say "de" (Dutch for "the") aloud, quickly (at a rate of two to three per second), and without stopping while they were solving the sums. This secondary task was meant to load the phonological loop. Before the start of this condition, the experimenter instructed the participants that it was very important not to stop articulating while solving the sums. Moreover, the experimenter carefully observed whether the secondary task was performed as instructed. Because of this, there is no reason at all to assume that this secondary task was not performed accurately. The RIG task (Vandierendonck et al., 1998) was used as a third condition: The participants were asked to tap an unpredictable "rhythm" on the zero key of the numeric keypad while they were solving the sums. They were instructed to use their left index finger and were told that the "rhythm" had to be as random and unpredictable as possible. This task was meant to load the central executive. In order to get a measure of randomness, the tap sequences of the participants were registered. The sequence of the conditions was fully counterbalanced, and the participants were permitted a 3-min rest period between the conditions.

At the start of the experiment, the participants solved 16 random practice trials (eight true sums, four with a split of ± 1, four with a split of ± 9) in order to familiarize themselves with the apparatus, the procedure, the stimulus display, and the response key assignment. After each practice trial, the participants received feedback. Afterward, no more feedback was provided. Each series started with a fixation point (a !) in the middle of the screen that remained for 500 msec. In the experimental conditions, the participants were first required to exercise the concurrent task until they felt comfortable with it. They were told that is was important not to stop performing the secondary task while they were solving the sums. Experiments 1 and 2 were administered in one experimental session. The sequence of both experiments was counterbalanced. In order to investigate whether a tradeoff between primary and secondary tasks appeared, the participants were asked to perform the RIG task alone for 5 min. Ten participants did so before both experiments, 10 between the experiments, and 10 after both experiments. The participants were tested individually in a quiet room. Each experimental session lasted approximately 50–70 min.

Results

A 3 (load, CON, AS, or RIG) $\times$ 2 (size, small and large) $\times$ 3 (sum, true, split = ± 1 or split = ± 9) within-subjects design analysis of variance (ANOVA) was used to analyze the primary task data. The sum of the *small* problems was less than 10; the *large* problems had a correct answer equal to or larger than 10 (this distinction was based on Hamann & Ashcraft, 1985, and LeFevre, Sadesky, & Bisanz, 1996). All reported effects are significant at $p < .05$ at least, unless otherwise stated.

The analyses of the RIG task, which will be described later, revealed that the patterns of 4 participants deviated from randomness and differed greatly from the patterns of the other participants. These participants were excluded from all further analyses.

Latencies. Only the latencies of the correctly solved sums were included in the analysis. Mean latencies, standard deviations, and accuracy are displayed in Table 1.

There were significant main effects of load [$F(2,50) = 21.25$], of size [$F(1,25) = 72.21$], and of split [$F(2,50) = 72.80$]. Articulatory suppression had no effect on true sums [$F(1,25) < 1$] or on sums with a small [$F(1,25) < 1$] or a large [$F(1,25) = 2.37, p > .10$] split. On the contrary, the RIG task yielded latencies significantly higher than those in the CON condition for the true sums [$F(1,25) = 36.04$] and for the sums with small and large splits [$F(1,25) = 5.87$ and $F(1,25) = 29.03$, respectively].

The size $\times$ split interaction was the only interaction found to be significant [$F(2,50) = 16.20$]. Further analyses revealed that the difference between the small and the large problems was significant for the true problems (Tukey's HSD, $p < .001$) and for the small-split problems (Tukey's HSD, $p < .001$), but not for the large split-problems (Tukey's HSD, $p > .50$).

Accuracy. As can be seen in Table 1, accuracy was very high. Consequently, these data are not very sensitive, and differences in accuracy must be interpreted with caution. These analyses revealed more or less the same pattern: The main effects of size and of split and their interaction were significant [$F(1,25) = 12.04$, $F(2,50) = 17.48$, and $F(2,50) = 4.63$, respectively]. Also, the three-way load $\times$ size $\times$ split interaction was significant [$F(4,100) = 4.53$]. However, because this interaction (1) was not found in the latencies and (2) was not replicated in Experiment 2 (see below), we do not interpret this finding as being robust.

Randomness analyses. A global overview, rather than a detailed description, of these analyses is given. We refer the interested reader to Vandierendonck (2000a; Vandierendonck et al., 1998); the detailed analyses are available from the authors. In essence, random time intervals can be converted into a series of binary events by dividing the complete time course into a series of fixed intervals that either contain or do not contain a tap. By means of the appropriate statistics on the probability of the taps, the tendency to deviate from randomness (i.e., to alternate or to perseverate) can be estimated. These statistics and the analyses revealed that the RIG task was performed equally well in the experimental condition as in the single secondary task CON condition, indicating that the participants complied with task instructions and that no tradeoffs between the primary and the secondary tasks occurred.

Discussion

The main purpose of Experiment 1 was to further explore the effect of articulatory suppression on the verification of true sums. This secondary task was found to have no effect, suggesting that the phonological loop is not involved. On the contrary, the RIG task hindered the verification of the sums, extending the conclusions of Lemaire et al. (1996) and of De Rammelaere et al. (1999), who argued for a crucial contribution of the central executive to the verification of simple arithmetic sums.

Although this experiment was not explicitly designed to test this, we obtained an interesting finding about the split effect (i.e., the finding that a false answer that is far from the correct answer, such as 8 + 4 = 21, is rejected faster than a false answer that is closer to the correct answer, such as 8 + 4 = 13). This effect has been reported frequently (e.g., Ashcraft & Battaglia, 1978; De Rammelaere et al., 1999; Stazyk, Ashcraft, & Hamann, 1982; Zbrodoff & Logan, 1990), but at this point it is not possible to choose between two possible explanations. According to a first explanation, the correct answer is retrieved and compared with the presented answer. Because it takes longer to compare two *close* than two *far* numbers (Moyer & Landauer, 1967), *far* false answers are rejected faster than *close* false answers. The second explanation, on the contrary, does not assume that the correct answer is always activated. Instead, participants use two different strategies (for strategy use in arithmetic, see, e.g., Lemaire & Fayol, 1995; Lemaire & Reder, 1999; Lemaire & Siegler, 1995). The *retrieval* strategy is used to verify *close* false answers: Participants first retrieve the correct solution, compare the retrieved solution with the presented answer, and state their response. The *plausibility* strategy is used to verify *far* false answers: The whole verification process is not run to completion but is short-circuited by a *fast-no* decision, given the implausibility of the proposed answer. The important difference between the two explanations is that, with large splits, the correct answer is activated according to the first explanation, but not according to the second explanation. Interestingly, our data are more in line with the second explanation: The problem size effect (i.e., larger problems are more difficult than smaller problems) appeared for the true and the small-split problems, but not for the large-split problems, suggesting that the correct answer was not activated for the latter problems.

EXPERIMENT 2

In Experiment 2, we wanted to test whether the findings of Experiment 1 are also valid for multiplication problems. More specifically, we tested three hypotheses. (1) Articulatory suppression will not interfere with simple multiplication problems, even not with true ones. (2) RIG will hinder true, small-split, and large-split problems. (3) We will obtain a size × split interaction, so that a problem size effect appears for true and small-split problems but disappears for large-split problems.

Method

Participants. The same 30 subjects as those in Experiment 1 also participated in Experiment 2.

Stimuli, Procedure, and Design. The stimuli, procedure, and design were identical to those in Experiment 1. The only difference was that multiplication was studied instead of addition. Thus, there were again three kinds of problems: true, small split, and large split. However, in Experiment 2, we worked with a relative instead of an absolute split (following Stazyk et al., 1982). This modification was introduced because the range for multiplication is much larger than that for addition: from 4 (2 × 2) to 81 (9 × 9) in multiplication, but "only" from 4 (2 + 2) to 18 (9 + 9) in addition. Because of this, it is not certain that the split of ±9 in Experiment 1 is sufficiently large.

To get false answers, the correct answer was multiplied by 0.1 (large negative split), 0.9 (small negative split), 1.1 (small positive split), or 1.9 (large positive split). These splits were chosen in order to get a large difference between small and large splits (see, e.g., Stazyk et al., 1982). Then, this number was rounded to the nearest integer that satisfied three conditions. (1) If the correct answer was even (odd), the false answer had to be odd (even). (2) The false answer could never be a multiple of a or b. (3) The false answer could never be the sum of $a + b$. To illustrate, consider an example. Suppose the problem 9 × 3 belongs to the group of positive splits. In order to get a false answer with a small positive split, the correct answer is multiplied by 1.1 (27 × 1.1 = 29.7). The closest *even* number to 29.7 is 30. However, 30 is not possible, since 30 is a multiple of 3. Because of this, the second smallest number, 28, is in this example chosen as a false answer. The number 52 is chosen as a false answer with a large positive split, since this number is the closest even number to 51.3 (27 × 1.9 = 51.3).

Results and Discussion

The same data-analytic strategy as that in Experiment 1 was used. In order to keep the two experiments comparable, the same problems as those in Experiment 1 were assigned to the small and the large groups. The analyses of the RIG task revealed that the pattern of the same 4 participants deviated from randomness and differed greatly from the patterns of the other participants. As in Experiment 1, these participants were excluded from all further analyses.

Latencies. Only the latencies of the correctly solved products were included in the analysis. Mean latencies, standard deviations, and accuracy for each kind of product in each condition are shown in Table 2.

There were significant main effects of load [$F(2,50) = 13.40$], of size [$F(1,25) = 61.25$], and of split [$F(2,50) = 40.50$]. Articulatory suppression had no effect on true, small-split or large-split problems [all Fs$(1,25) < 1$], confirming Hypothesis 1. The RIG task, on the contrary, yielded latencies significantly higher than those in the CON condition for the true problems [$F(1,25) = 14.68$] and the problems with a small [$F(1,25) = 16.55$] and a large [$F(1,25) = 10.92$] split, confirming Hypothesis 2.

Again, the size × split interaction was the only significant interaction [$F(2,50) = 22.32$]. Further analyses

Table 2
Experiment 2: Mean Latencies (RTs, in Milliseconds), Standard Deviations, and Accuracy for the Combinations of Load, Size, and Split

		CON			AS			RIG		
Size	Split	RT	*SD*	AC	RT	*SD*	AC	RT	*SD*	AC
Small	true	880	217	.95	899	180	.94	1,037	245	.93
	small	994	218	.97	959	168	.92	1,174	305	.93
	large	1,012	272	.95	1,043	268	.95	1,187	381	.95
Large	true	1,031	233	.96	1,017	218	.93	1,160	251	.93
	small	1,145	234	.92	1,197	306	.90	1,305	367	.93
	large	994	248	.96	981	159	.97	1,144	270	.97

Note—CON, control; AS, articulatory suppression; RIG, random time interval generation; AC, accuracy.

showed that the difference between the small and the large problems was significant for the true problems (Tukey's HSD, $p < .001$) and for the small-split problems (Tukey's HSD, $p < .001$), but not for large-split problems (Tukey's HSD, $p > .50$), confirming Hypothesis 3.

Accuracy. The accuracy data revealed only a significant main effect of split [$F(2,50) = 6.88$] and a significant size × split interaction [$F(2,50) = 4.09$].

Randomness analyses. The sequences of random taps were analyzed in the same way as in Experiment 1. Again, these analyses suggested that the participants complied with task instructions and indicated no primary–secondary task tradeoffs.

GENERAL DISCUSSION

In the present study, the participants verified single one-digit addition (Experiment 1) and multiplication (Experiment 2) problems in a CON condition, an AS condition, and a condition with the RIG task (Vandierendonck et al., 1998). The problems were true or had a small or a large split. Three general conclusions can be drawn. The first two conclusions concern the contributions of the phonological loop and the central executive, respectively; the third conclusion deals with different arithmetic strategies to solve small- and large-split problems.

Our data suggest that the phonological loop is not involved in the verification of false arithmetic problems or in the verification of true problems (e.g., 8 + 4 = 12 or 7 × 3 = 21). The latter finding is not in agreement with models that assume that basic arithmetic facts are stored in a language-dependent verbal form, like Dehaene's triple-code model (Dehaene, 1992; Dehaene & Cohen, 1995). In fact, other studies also have suggested that this assumption of the model might be wrong. Pesenti, Thioux, Seron, and De Volder (2000), for instance, found that none of the cerebral language areas usually involved in verbal processes were activated during arithmetic fact retrieval. Brysbaert et al. (1998) found some important language differences between Dutch- and French-speaking participants in the latencies of addition problems but discovered that these had been caused by output requirements, rather than by language-specific addition processes. Interestingly, their conclusions were very similar to those of Noël et al. (1997), who found interactions between reading processes and multiplication solutions but were able to show that these interactions did not occur at the stage of arithmetic fact retrieval, but at the later, output stage (see also Campbell, 1998; Noël et al., 1998). In an attempt to find the underlying cause of children's arithmetical difficulties, Bull and Johnston (1997) and Bull, Johnston, and Roy (1999) concluded that "specifically the functioning of the articulatory loop did not represent a fundamental deficit for children of low mathematical ability" (Bull et al., 1999, p. 421). McLean and Hitch (1999) showed that children with poor arithmetic skills had, relative to age-matched control subjects, normal phonological working memory. In fact, our results are also in agreement with the case study reported by Butterworth, Cipolotti, and Warrington (1996), who concluded that although their patient, M.R.F., "had an impaired articulatory loop, his arithmetical skills remained satisfactory" (p. 261).

The second conclusion concerns the contribution of the central executive. Once more, it was found that a secondary task taxing executive processes interfered with the verification of simple arithmetic problems. Besides the already-mentioned studies (Ashcraft et al., 1992; De Rammelaere et al., 1999; Hitch, 1978; Lemaire et al., 1996; Logie et al., 1994), other studies also have shown that executive processes might be crucial in arithmetic. Bull et al. (1999) revealed that children of high and low mathematical ability differed especially on measures of the central executive. McLean and Hitch (1999) demonstrated that children with poor arithmetic skills were impaired particularly on executive processing. Klein and Bisanz (2000) found that working memory constraints especially limit the arithmetical performance of preschool children. Our results suggest that the central executive has a general effect on processing (i.e., no significant load × size interaction) but do not tell us which aspects of the central executive are important to arithmetic. We speculate, however, that the load × size interaction would be significant when more difficult problems, such as 16 + 49 or 17 × 6, are included. Note that the effect of RIG cannot be due to a tradeoff between primary and secondary tasks, because the analyses revealed that the participants tapped equally randomly in both conditions. Neither can it be argued that the effect of RIG is due simply to the fact that two tasks had to be executed simultaneously, because in the articulatory suppression condition also, two tasks had to be executed simultaneously. Therefore, we conclude that the central executive is involved in the verification of simple mental arithmetic problems, but we do not know (yet) which executive process(es) do contribute or in what manner.[1] In order to find an answer to this question, one needs to fractionate the central executive into different executive processes, which is only a new, although very promising, research area (e.g., Miyake,

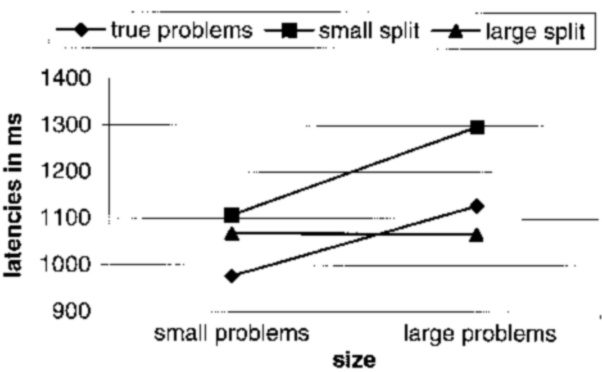

Figure 1. Mean latencies (in milliseconds) for the problems (averaged over addition and multiplication) as a function of size and of split.

Friedman, Emerson, Witzki, & Howerter, 2000). Furthermore, the development of valid instruments or secondary tasks to measure or to load specifically these different executive processes is not without problems (e.g., Rabbitt, 1997).

The problem size × split interaction is the third important finding of the present study. Although this interaction was explored earlier by a number of researchers (e.g., Ashcraft & Stazyk, 1981; Campbell, 1987; Geary, Widaman, & Little, 1986; Stazyk et al., 1982; Zbrodoff & Logan, 1990), none of these studies tested explicitly whether a problem size effect occurred with large splits or interpreted their results in the light of (non)retrieval of the correct answer. We obtained evidence for the fact that the split effect is due to the use of two different strategies and is not simply the consequence of the fact that the comparison of the correct (retrieved) answer goes faster with *far* than with *close* false answers.

Figure 1 shows that the line connecting the small and the large problems was almost exactly parallel for the true and the small-split problems. In contrast, the problem size effect completely disappeared for the large splits. These results are more in line with an explanation of the split effect in terms of two different strategies than in terms of one common comparison mechanism. For small-split problems, the retrieval strategy is used: Participants first retrieve the correct answer and compare this number with the presented one. Large-split problems, on the contrary, are solved by means of the plausibility strategy, in which a *fast-no* is immediately decided upon, instead of retrieving the correct answer. Dehaene's triple-code model can handle the outcomes concerning the use of two different strategies perfectly. Indeed, the model explicitly states that retrieving basic arithmetic facts and making approximations use different representations and take place in different brain areas. Moreover, our results fit with the case study of Dehaene and Cohen (1991), who described a patient who was able to say that 2 + 2 = 9 is wrong but could not reject 2 + 2 = 5. In other words, the approximation of this patient was preserved, which led Dehaene and Cohen (1991) to conclude that there are two mental calculation systems.

REFERENCES

Ashcraft, M. H., & Battaglia, J. (1978). Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. *Journal of Experimental Psychology: Human Learning & Memory*, **4**, 527-538.

Ashcraft, M. H., Donley, R. D., Halas, M. A., & Vakali, H. (1992). Working memory, automaticity, and problem difficulty. In J. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 301-329). Amsterdam: Elsevier.

Ashcraft, M. H., & Stazyk, E. H. (1981). Mental addition: A test of three verification models. *Memory & Cognition*, **9**, 185-196.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-89). New York: Academic Press.

Brysbaert, M., Fias, W., & Noël, M. P. (1998). The Whorfian hypothesis and numerical cognition: Is "twenty-four" processed in the same way as "four-and-twenty"? *Cognition*, **66**, 51-77.

Bull, R., & Johnston, R. S. (1997). Children's arithmetical difficulties: Contributions from processing speed, item identification, and short-term memory. *Journal of Experimental Child Psychology*, **65**, 1-24.

Bull, R., Johnston, R. S., & Roy, J. A. (1999). Exploring the roles of the visual-spatial sketch pad and central executive in children's arithmetical skills: Views from cognition and developmental neuropsychology. *Developmental Neuropsychology*, **15**, 421-442.

Butterworth, B., Cipolotti, L., & Warrington, E. K. (1996). Short-term memory impairment and arithmetical ability. *Quarterly Journal of Experimental Psychology*, **49A**, 251-262.

Campbell, J. I. D. (1987). Production, verification, and priming of multiplication facts. *Memory & Cognition*, **15**, 349-364.

Campbell, J. I. D. (1998). Linguistic influences in cognitive arithmetic: Comment on Noël, Fias and Brysbaert (1997). *Cognition*, **67**, 353-364.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, **44**, 1-42.

Dehaene, S., & Cohen, L. (1991). Two mental calculations systems: A case study of severe acalculia with preserved approximation. *Neuropsychologia*, **29**, 1045-1074.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition*, **1**, 83-120.

De Rammelaere, S., Stuyven, E., & Vandierendonck, A. (1999). The contribution of working memory resources in the verification of simple mental arithmetic sums. *Psychological Research*, **62**, 72-77.

Geary, D. C., Widaman, K. F., & Little, T. D. (1986). Cognitive addition and multiplication: Evidence for a single memory network. *Memory & Cognition*, **14**, 478-487.

Hamann, M. S., & Ashcraft, M. H. (1985). Simple and complex mental addition across development. *Journal of Experimental Child Psychology*, **40**, 49-72.

Hitch, G. J. (1978). The role of short-term working memory in mental arithmetic. *Cognitive Psychology*, **10**, 302-323.

Klein, J. S., & Bisanz, J. (2000). Preschoolers doing arithmetic: The concepts are willing but the working memory is weak. *Canadian Journal of Experimental Psychology*, **54**, 105-116.

LeFevre, J., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 216-230.

Lemaire, P., Abdi, H., & Fayol, M. (1996). The role of working memory resources in simple cognitive arithmetic. *European Journal of Cognitive Psychology*, **8**, 73-103.

Lemaire, P., & Fayol, M. (1995). When plausibility judgments supersede fact retrieval: The example of the odd–even effect on product verification. *Memory & Cognition*, **23**, 34-48.

Lemaire, P., & Reder, L. (1999). What affects strategy selection in arithmetic? The example of parity and five effects on product verification. *Memory & Cognition*, **27**, 364-382.

Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change:

Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General*, **124**, 83-97.

Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition*, **22**, 395-410.

McLean, J. F., & Hitch, G. J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology*, **74**, 240-260.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable approach. *Cognitive Psychology*, **41**, 49-100.

Moyer, R. S., & Landauer, T. K. (1967). The time required for judgements of numerical inequality. *Nature*, **215**, 1519-1520.

Noël, M. P., Fias, W., & Brysbaert, M. (1997). About the influence of the presentation format on arithmetical-fact retrieval processes. *Cognition*, **63**, 335-374.

Noël, M. P., Robert, A., & Brysbaert, M. (1998). Does language really matter when doing arithmetic? Reply to Campbell (1998). *Cognition*, **67**, 365-373.

Pesenti, M., Thioux, M., Seron, X., & De Volder, A. (2000). Neuroanatomical substrates of Arabic number processing, numerical comparison and simple addition. *Journal of Cognitive Neuroscience*, **12**, 461-479.

Rabbitt, P. (1997). *Methodology of frontal and executive function.* Hove, U.K.: Psychology Press.

Stazyk, E. H., Ashcraft, M. H., & Hamann, M. S. (1982). A network approach to mental multiplication. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **8**, 320-335.

Vandierendonck, A. (2000a). Analyzing human random time generation behavior: A methodology and a computer program. *Behavior Research Methods, Instruments, & Computers*, **32**, 555-565.

Vandierendonck, A. (2000b). Bias and processing capacity in the generation of random time intervals. *Cognitive Science Quarterly*, **1**, 205-233.

Vandierendonck, A., De Vooght, G., & Van der Goten, K. (1998). Does random time interval generation interfere with working memory executive functions? *European Journal of Cognitive Psychology*, **10**, 413-442.

Zbrodoff, N. J., & Logan, G. D. (1990). On the relation between production and verification tasks in the psychology of simple arithmetic. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 83-97.

NOTE

1. Recently, Vandierendonck (2000b) proposed a model that describes which executive functions could be involved in the RIG task. The basic idea is that a decision process listens to the pulses of a timing mechanism and acts on it. Another executive function consists of monitoring the motor output and warning the decision process when deviations from the notion of randomness are perceived. The core of the model thus consists of two executive functions: decision making and monitoring. However, planning of sequences of taps and the inhibition of habitual responses (known rhythms) could also be involved. Future research is required to clarify which of these executive functions interferes with the verification of arithmetic problems.

(Manuscript received February 2, 2000;
revision accepted for publication September 13, 2000.)