Strategy use in mental subtraction determines central executive load

LOEL N. TRONSKY AND MELISSA MCMANUS
Albertus Magnus College

ERIC C. ANDERSON
Hampshire College

A dual task method was used to examine the relationship between strategy use and working memory load during subtraction problem solving. Undergraduates mentally solved subtraction problems alone and while performing secondary tasks that involved the central executive of working memory. Analyses revealed that a central executive task involving response selection and input monitoring (CRT-R task) interfered more with subtraction problem solving than a task that involved only input monitoring (SRT-R task). Additional analyses showed that the CRT-R task interfered more when participants used a nonretrieval (counting) strategy than a retrieval strategy. These findings suggest that the response selection subcomponent of the central executive is involved during both retrieval-based and non-retrieval-based simple subtraction problem solving but is involved more during the latter.

Recently researchers have documented that use of nonretrieval procedures, such as counting, to solve simple arithmetic problems (e.g., $5 + 9$, $13 - 8$, $7 \times 4$) is quite common; depending on the sample, nonretrieval procedures are used on 30% of problems or more (Campbell & Xue, 2001; Hecht, 2002; LeFevre et al., 1996; Seyler, Kirk, & Ashcraft, 2003; Tronsky & Shneyer, 2004). Before the aforementioned studies it was often implicitly or explicitly assumed that adults use retrieval from long-term memory to solve most, if not all, basic arithmetic problems (Ashcraft & Battaglia, 1978; Ashcraft, Donley, Halas, & Vakali, 1992; Lemaire, Abdi, & Fayol, 1996). Little is known about the effect of differential strategy use on some well-documented empirical effects in mental arithmetic. For example, only a few studies have examined the relationship between adults’ use of strategies and working memory (WM) involvement in arithmetic processing (Hecht, 2002; Seyler et al., 2003; Tronsky & Shneyer, 2004). The main goal of the present investigation was to further our understanding of the relationship between strategy use in mental arithmetic and a specific subcomponent of WM, the central executive. Although an exhaus-
tive review of the literature on adult strategy use and WM involvement in mental arithmetic is beyond the scope of this article, an introduction to these topics is warranted (see LeFevre, Smith-Chant, Hiscock, Daley, & Morris, 2003, for a review of adult strategy use in arithmetic and DeStefano & LeFevre, 2004, and Tronsky & Royer, 2003, for a review of WM involvement in arithmetic).

Most WM investigations of arithmetic have used Baddeley’s (1996, 2001) WM model. In this model, WM is described as a four-part, limited capacity system that allows one to simultaneously store and process information. More specifically, WM is composed of a multifunctional attentional system called the central executive and three subsidiary or slave components: the phonological loop, the visuospatial sketchpad, and the episodic buffer. The phonological loop is used to store and rehearse auditory and verbal information, and the visuospatial sketchpad is used to store and rehearse visual and spatial information. A recently added component, the episodic buffer, remains largely untested, but one of its most important functions might be to combine information from the other slave systems and episodic long-term memory.

The central executive has many functions and can be fractionated into correlated but distinct subcomponents that are responsible for processes such as inhibition (suppressing task irrelevant information), strategy shifting (changing problem-solving strategies), memory updating (modifying current memory representations to accommodate new input), input monitoring (scanning incoming information for elements relevant to currently performed tasks), and response selection (Baddeley, 1996; Miyake et al., 2000; Szmalec, Vandierendonck, & Kemps, 2005; Vandierendonck, de Vooght, & Van der Goten, 1998). Determining WM involvement in cognitive tasks often involves using dual tasks. In these methods participants perform a primary task, such as arithmetic, while simultaneously performing a secondary task that loads a specific component of WM, such as having participants constantly repeat a word or nonword to load the phonological loop. WM involvement is indicated by poorer performance on primary or secondary tasks during dual task conditions compared with primary and secondary task performance when the tasks are performed separately.

To date findings are somewhat mixed regarding the role the slave systems of WM play in simple mental arithmetic; some have shown that the phonological loop and visuospatial sketchpad are involved (Ashcraft et al., 1992; Hecht, 2002; Lee & Kang, 2002; Lemaire et al., 1996), whereas others have shown they are not (De Rammelaere, Stuyven, & Vandierendonck, 2001; De Rammelaere & Vandierendonck, 2001; Seitz & Schumann-Hengsteler, 2000, 2002). In contrast, most studies have shown that the central executive is loaded during simple mental arithmetic (De Rammelaere et
al., 2001; De Rammelaere & Vandierendonck, 2001; Deschuyteneer & Vandierendonck, 2005a, 2005b; Lemaire et al., 1996; Seitz & Schumann-Hengsteler, 2000, 2002; see Rusconi, Galfano, Speriani, & Umilta, 2004, for a different argument). More recently, in response to findings that the central executive can be fractionated, researchers have begun to examine which subcomponents of the central executive are involved in arithmetic problem solving (Deschuyteneer & Vandierendonck, 2005a, 2005b; Deschuyteneer, Vandierendonck, & Muyllaert, 2006).

For example, Deschuyteneer and Vandierendonck (2005a) had participants solve simple addition problems while performing dual tasks that used the central executive functions of input monitoring or response selection. For the input monitoring tasks, participants listened to a tone presented at fixed time intervals (simple reaction time task, fixed; SRT-F) or random time intervals (simple reaction time task, random; SRT-R) and pressed the “0” key on the numeric pad in response to each tone. For the response selection tasks participants listened to two different tones (one low frequency, one high frequency) presented at fixed time intervals (choice reaction time task, fixed; CRT-F) or random time intervals (choice reaction time task, random; CRT-R) and pressed the “1” key on the numeric pad in response to the low tone and the “4” key in response to the high tone. Across four experiments Deschuyteneer and Vandierendonck found that addition response times (RTs) and errors increased from single to dual task conditions, as did secondary task tapping RTs and errors. They also found that the central executive component of input monitoring was minimally involved in addition problem solving; the SRT-F task (which entails minimal input monitoring) disrupted arithmetic RTs and accuracies as much as the SRT-R task (which entails greater input monitoring). Their Experiment 2 yielded similar results; the two choice reaction time tasks (CRT-F and CRT-R) that differed only in input monitoring interfered with addition RTs approximately equally. In contrast, response selection affected addition problem solving; the CRT-F task, which entailed response selection, affected addition RTs and accuracies more than the SRT-F task, which did not entail response selection. The comparison of CRT-R and SRT-R interference yielded parallel findings. Deschuyteneer and Vandierendonck (2005b) replicated these results with mental multiplication as the primary task.

Something that clouds the aforementioned arithmetic and WM findings reviewed thus far is the recent documentation that adults often use strategies other than retrieval to solve basic arithmetic problems. Three of the most common nonretrieval strategies used by adults are the derived fact strategy (also called decomposition), counting strategy, and referencing a complementary operation strategy (Hecht, 2002; LeFevre et al., 1996; Tronsky & Shneyer, 2004). The derived fact strategy involves using
a problem whose answer can be easily retrieved from long-term memory to help solve a problem whose answer cannot be retrieved from long-term memory. For example, some adults solve the problem $6 + 7$ by retrieving the answer to $6 + 6$ and adding one to it (i.e., $6 + 6 = 12$, $12 + 1 = 13$). Adults also use counting strategies somewhat frequently (Hecht, 2002; Seyler et al., 2003). The most common counting method in addition is the min counting strategy, which involves implicitly (or explicitly) counting up from the larger number in the problem (e.g., $3 + 5$ could be solved by implicitly counting “$6, 7, 8$”). The third type of strategy that adults often use for subtraction and division problems is referencing a complementary operation (Seyler et al., 2003; Tronsky & Shneyer, 2004). For example, to solve the problem $13 - 6$, adults sometimes report solving the corresponding addition problem $6 + ? = 13$. Other strategies are used (e.g., the nines rule in multiplication, where for $9 \times a$, $a - 1$ is the tens digit in the answer and $9 - (a - 1)$ is the ones digit) but are much less common than the three types of strategies already described.

In light of the aforementioned strategy findings, two recent investigations have been conducted to examine WM involvement across strategy use. Seyler et al. (2003) showed in their Experiment 3 that nonretrieval strategy use increased dramatically for larger subtraction problems (problems that had minuends of 11 or greater). In their Experiment 4, adults solved simple subtraction problems in several conditions, alone or while trying to remember a two-, four-, or six-letter consonant string. Letter recall performance was significantly poorer in the load conditions than in the no load condition, particularly for the six-letter load, and errors increased significantly more across load conditions for larger problems. Combining the results of these two experiments suggests that WM is more heavily loaded when people use nonretrieval strategies to solve simple subtraction problems.

In another study, Hecht (2002) presented addition verification problems (e.g., $6 + 7 = 14$) and had participants determine whether the equations were true or false by pressing the appropriate computer key. Participants solved these problems in three different conditions: addition alone, addition with articulatory suppression (repeating a letter of the alphabet, designed to load the phonological loop), and addition with random letter generation (producing a random sequence of letters, designed to load the central executive). Hecht also had participants report the strategies they used to solve each problem in each condition.

Analyses yielded three important findings. First, strategy use was consistent across the three different memory load conditions for true equations and with the exception of one strategy was consistent for false equations. This showed that adding a WM load did not appreciably change the strategies participants used. Second, both types of WM load caused participants’
counting processes to slow and resulted in poorer secondary task performance. In contrast, neither type of WM load disrupted core retrieval processes; retrieval rates remained constant across load conditions, and retrieval rates did not interact with secondary task performance.

EXPERIMENT

The purpose of this experiment was to further our understanding of the strategy use–WM relationship in simple mental arithmetic. This was accomplished using two tasks that load the input monitoring and response selection components of the central executive. Use of these tasks builds on the investigations of Deschuyteneer and Vandierendonck (2005a, 2005b), Hecht (2002), and Seyler et al. (2003). First, SRT-R and CRT-R are pure tasks that load the central executive without loading the slave systems of WM (Szmalec et al., 2005; Vandierendonck et al., 1998). These secondary tasks will help clarify the results of studies that have found WM involvement in arithmetic using tasks that load more than one part of the WM system; both Hecht and Seyler et al. used central executive tasks that may have loaded the phonological loop as well (e.g., random letter generation). Also, to date no one has examined the role of response selection in retrieving subtraction answers from long-term memory, and no one has examined whether this central executive function is more involved during non–retrieval-based or retrieval-based problem solving for any arithmetic operation.

In our experiment, participants solved subtraction problems alone and while responding to auditory tones in the SRT-R and CRT-R tasks from Deschuyteneer and Vandierendonck (2005a, 2005b). In both single and dual task conditions participants reported the strategies they used to solve each subtraction problem. Deschuyteneer and Vandierendonck noted that response selection may be involved in arithmetic answer retrieval for the following reason. For adults it is well documented that the presentation of arithmetic problems automatically activates candidate answers from long-term memory (Galfano, Rusconi, & Umiltà, 2003; Rusconi et al., 2004). This activation includes both the correct answer and incorrect but related answers. The selection of the correct answer from this set of potential answers may be the source of the response selection interference in the studies conducted by Deschuyteneer and Vandierendonck. Based on these findings and arguments, we predict that the CRT-R dual task will interfere with subtraction retrieval speed or accuracy more than the SRT-R task will.

Turning to non–retrieval-based subtraction problem solving, several decision components may be involved. For example, a counting strategy
might be used instead of retrieval. Two different types of counting strategies might be used. One may count up (e.g., 9 – 6 solved by counting “7, 8, 9, the answer is 3”) or count down (e.g., 9 – 3, solved by counting down “8, 7, 6, the answer is 6”) to solve a problem, and the process of making that decision may use central executive resources. Another decision may be needed toward the end of the count, more specifically whether the requisite number of counts has been performed (whether the end of the count has been reached). Given these arguments, we predict that the CRT-R task will interfere more with both non-retrieval-based and retrieval-based subtraction problem solving than the SRT-R task because of its response selection component (which includes a decisional stage about which response should be produced; Szmalec et al., 2005). We also predict that, because of the multiple decision components involved in nonretrieval problem solving, CRT-R interference will be greater during non-retrieval-based than retrieval-based problem solving.

METHOD

Participants

Twenty-three participants from a small New England college, 18 women and 5 men with an average age of 20.1 years, volunteered to participate and received extra credit toward their psychology classes for completing the experiment. All participants reported English as their first language, none reported a diagnosed reading disability, and one participant reported a diagnosed mathematical disability. The results did not vary when this participant’s data were included; therefore, the participant’s data were retained in the analyses reported.

Materials

Visual stimuli were presented on a PowerMac G4 computer using SuperLab (2.0) and appeared on the screen in black font against a white background. A PC running SuperLab (2.0.2) presented auditory stimuli through headphones, which were used to control the volume of the stimuli and to minimize ambient noise. Answers to the visual-based tasks were spoken into a microphone connected to the computer, and responses to the auditory stimuli were made via keyboard presses. Stimuli and tasks were presented in a different random order for each participant.

Computer-administered tasks

Number naming task. Stimuli for this task were Arabic numbers that were answers to simple subtraction or multiplication (from another experiment) problems. Each of the 32 multiplication answers was presented once, and each of the 8 subtraction answers was presented four times, for a total of 64 stimuli. The task began with a 1,000-ms blank screen followed by one of the Arabic numbers. After participants named the number, the RT for that stimulus was recorded, and the screen went
blank for 1,000 ms. The sequence was repeated until all 64 stimuli had been presented. This task was included as a comparison and control task for the problems that were solved by retrieval in the subtraction tasks. Previously, tone and arithmetic tasks have been presented alone as controls (Deschuyteneer & Vandierendonck, 2005a, 2005b). Using the number naming task with the tone tasks rather than the tone tasks alone is more appropriate as a control because it differs from the arithmetic dual task conditions on fewer dimensions. Both the number naming dual task and arithmetic dual task conditions entail the coordination of two tasks and entail a verbal production component; presenting the tone tasks by themselves does not.

The number naming with SRT-R and CRT-R dual tasks also are more appropriate control tasks for subtraction retrieval in particular because both tasks include an encoding phase (entering the number or numbers into memory), a retrieval of answer or name phase, and a phonological preparation and output phase (preparing to speak and then speaking the answer into the microphone). If significant response selection resources are needed when participants are retrieving answers to subtraction problems, RTs should increase more across memory load conditions during the retrieval of subtraction answers than during retrieval of number names. If not, it is evidence that processes peripheral to retrieving subtraction answers (e.g., problem encoding) may also use attentional resources.

**Subtraction task.** In this task, the 64 simple subtraction problems with answers greater than 1 (4 – 2 through 18 – 9) were presented. The task began with a 1,000-ms blank screen followed by a subtraction problem. After speaking their answers into the microphone, participants were immediately prompted by the words “Report Strategy” to explain how they had solved the problem. Upon completion of the verbal report, the experimenter pressed the spacebar on the keyboard, and after a 1,000-ms blank screen the sequence of events was repeated until all problems had been presented. Problems were presented in different random order to each participant in all conditions. A cassette recorder was used to document participants’ strategy reports, and after the experiment, reports were coded into strategy categories.

**SRT-R.** In this task participants responded to tones that were presented at random time intervals. They were 262-Hz tones (C1 note) 70 ms in duration that were presented either 900 or 1,500 ms after the start of the previous tone. Selection of the time interval for each tone presentation was random throughout the task. Participants responded by pressing the “1” key on the numeric keypad of the computer each time they heard a tone.

**CRT-R.** This task required participants to respond to two tones presented at random time intervals. The low tone was a 262-Hz tone (C1 note), and the high tone was a 524-Hz tone (C2 = C1 + one octave), each presented for 70 ms. Tones were presented in random order either 900 or 1,500 ms after the start of the previously presented tone. Participants had to indicate which tone was presented, a low or high tone, by pressing the “1” or “2” key on the numeric keypad of the computer keyboard, respectively.

**Single and dual task conditions**

Each task—number naming, subtraction, SRT-R, and CRT-R—was presented as a single task. For the dual task conditions the subtraction and number naming
tasks were completed with participants speaking into a microphone held in their nondominant hand while they responded to tones from the SRT-R or CRT-R task using the index finger (SRT-R) or index finger and middle finger (CRT-R) of their dominant hand.

Strategy report instructions and training

Participants were trained on how to give retrospective verbal reports immediately before their first subtraction task. The use of verbal reports to determine the cognitive processes during various tasks, including arithmetic, has been scrutinized recently (Ericsson & Kirk, 2000, 2001; Kirk & Ashcraft, 2001; Robinson, 2001; Seyler et al., 2003). Researchers have documented that there are three major problems in using verbal reports to determine cognitive processes: validity, reactivity, and demand characteristics. Validity is the extent to which the reported processes reflect the actual processes that are used in the problem-solving task, reactivity is the extent to which asking participants to report their processes induces them to change those processes, and demand characteristics are aspects of an experimental procedure that may suggest to participants what types of reports or strategies the experimenter is expecting (for a review of these biases in the context of arithmetic strategy reports see Kirk & Ashcraft, 2001). Instruction and training procedures that minimize these concerns have been developed for, and tested using, mental arithmetic stimuli (Ericsson & Kirk, 2001; Seyler et al., 2003). These methods were modified slightly for the present investigation.

One of the authors began the training procedure by reading the instructions and then had participants solve nonarithmetic problems while concurrently reporting their solution processes (concurrent verbal report). On subsequent problems, participants were asked to recall the exact thoughts they had had while solving a problem after having solved it (retrospective verbal report); participants were told that they would be giving the latter type of report during the experiment. As in Seyler et al. (2003), problems that probably were solved automatically by retrieval from long-term memory (e.g., “What is the letter immediately after ‘A’?”) and those that probably entailed many steps (e.g., “What is the fourth letter after ‘H’?”) were included in the training session. The distinction between reporting exact thoughts (e.g., saying “I, J, K, L”—the desired report) and summarizing or describing thoughts (e.g., “I counted through the alphabet”) was emphasized, as the latter is a potentially biased report. Once participants were finished with the training they were given five computer-administered arithmetic problems (problems that were not used in the data collection part of the experiment) to begin using the retrospective strategy report procedure. Any remaining questions participants had were then answered. If either the participant or the experimenter was not satisfied that the participant understood the procedure, additional instructions and practice problems were provided.

Procedure

Participants were individually tested in two sessions that each lasted 45 to 60 min. All tasks included a standard set of instructions and five practice problems.
For each person, tasks were administered in a pseudo-random order. Session 1 was randomly chosen to be the SRT-R or CRT-R session. In the SRT-R session some single tasks and all the dual tasks involving SRT-R were presented, and in the CRT-R session some single tasks and all the dual tasks involving CRT-R were presented. This ensured that participants would not have to switch back and forth between the two tone tasks within sessions, reducing the possibility of confusion. Across the sessions number naming, subtraction, SRT-R, and CRT-R tasks were performed alone, and the number naming and subtraction tasks were performed simultaneously with each of the tone tasks, resulting in three levels of WM load (no-load control, with SRT-R, and with CRT-R). Each participant completed tasks in a different pseudo-random order within sessions.

RESULTS

Strategy coding

A pilot investigation using two strategy coders showed that reliability of the current strategy report procedure and coding is high (intercoder agreement ranged between 96% and 98%). Because of this high reliability, only the first author (who also was a coder in the pilot investigation) listened to the cassettes and categorized participants' strategy use. Retrieval and six nonretrieval strategies were identified for subtraction and are described here with examples.

The problems on which participants reported thinking only the answer or the problem and the answer (e.g., “I thought 15 – 7 is 8”) were coded as “retrieval.” “Derived fact” involved solving a problem by retrieving an answer to another problem and determining the answer on the basis of the known problem (e.g., “15 – 5 is 10, 7 – 5 is 2, and 10 – 2 is 8” in response to the problem 15 – 7). No distinction was made between addition-, subtraction-, and multiplication-derived fact strategies (e.g., some solved 15 – 7 saying, “7 + 3 is 10, 10 + 5 is 15, 3 + 5 is 8”; others said, “7 x 2 is 14, 14 + 1 is 15, 7 + 1 is 8”). A third category, “nines rule,” was coded when participants reported using a special rule to solve problems that involved nine (“I saw 9 and thought 1 + 5 is 6” in response to the problem 15 – 9). The fourth category, “addition” (addition reference), involved referring to a known addition problem (e.g., “7 + 8 is 15,” in response to 15 – 7). The fifth category was “counting,” where participants reported counting in their head (“I thought 6, 7, 8” when solving the problem 8 – 5). No distinction was made between counting up from the subtrahend or down from the minuend. A sixth category, “other strategy,” was reserved for idiosyncratic nonretrieval strategies (e.g., “I saw 5 sticks and 3 sticks and then took away the group of 3 leaving 2” to solve 5 – 3). The seventh category, “unknown,” included reports that were inaudible, indeterminate, or “I don’t know” responses.
Subtraction RT and accuracy analyses across WM load conditions

The percentage of problems for which each participant used each strategy in the three memory load conditions was calculated and averaged across the sample; these percentages are presented in Table 1. Range of use of each strategy across participants and the percentage of the sample who used each strategy are provided as well. Data from four participants were not used in the subtraction RT and error analyses: Two participants failed to return for the second session of the experiment, one was unable to follow the strategy report instructions, and strategy data for one participant were lost through experimenter error. Spoiled (11.5% of control, 8.7% of subtraction with SRT-R, and 15.5% of subtraction with CRT-R trials) and incorrect answer RT trials were removed.

As has been done in previous arithmetic investigations, RTs and accuracies were averaged across items (i.e., for each problem) rather than across participants (Ashcraft et al., 1992; Hecht, 2002). As Hecht and others noted, this method has several advantages, one of which is providing approximately equal numbers of items for comparison across conditions. In regression analyses it ensures that participant variability and problem variability are not confounded. Table 2 summarizes mean subtraction RTs and accuracies across strategies and WM load conditions using the aforementioned method. As can be seen from Table 1, only the retrieval and counting solution methods were used by at least half of the sample across each of the WM conditions and were used on the majority of problems. The remaining strategies (addition, derived fact, nines rule, and other) were used less frequently across participants and often for a much smaller subset of problems. For the analyses we planned to conduct it was important to have a fairly large number of data points to have a stable estimate of average item RTs (and error rates) and to have approximately equal numbers of item RTs and error rates for each strategy to enter in the analysis of variance (ANOVA). Therefore, we decided to conduct analyses on the only two strategies, retrieval and counting, that met the aforementioned criteria.

Initial analyses were conducted to determine WM load effects within each strategy before we moved on to comparisons across strategies. An ANOVA with the single within-participant factor load (control, SRT-R, CRT-R) was performed on the subtraction retrieval trial RTs, revealing a significant effect, $F(2, 126) = 6.67, p < .01$. This effect was due to significantly longer RTs in the CRT-R condition than in the control condition, 191 ms longer, $F(1, 63) = 10.39, p < .01$, and SRT-R condition, 197 ms longer, $F(1, 63) = 8.13, p < .01$. A similar analysis conducted for counting trial RTs also yielded a significant load effect, $F(2, 110) = 26.62, p < .001$. It was found that counting RTs in the CRT-R condition were significantly longer than
Table 1. Subtraction strategies used across working memory load conditions

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Mean use (%)</th>
<th>Percentage of sample</th>
<th>Range of use (%)</th>
<th>Mean use (%)</th>
<th>Percentage of sample</th>
<th>Range of use (%)</th>
<th>Mean use (%)</th>
<th>Percentage of sample</th>
<th>Range of use (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval</td>
<td>44</td>
<td>94</td>
<td>0–100</td>
<td>44</td>
<td>83</td>
<td>0–100</td>
<td>50</td>
<td>94</td>
<td>0–100</td>
</tr>
<tr>
<td>Counting</td>
<td>29</td>
<td>67</td>
<td>0–89</td>
<td>28</td>
<td>67</td>
<td>0–86</td>
<td>25</td>
<td>67</td>
<td>0–88</td>
</tr>
<tr>
<td>Addition</td>
<td>16</td>
<td>39</td>
<td>0–100</td>
<td>16</td>
<td>48</td>
<td>0–100</td>
<td>17</td>
<td>39</td>
<td>0–86</td>
</tr>
<tr>
<td>Derived fact</td>
<td>7</td>
<td>33</td>
<td>0–30</td>
<td>7</td>
<td>44</td>
<td>0–31</td>
<td>4</td>
<td>39</td>
<td>0–21</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>44</td>
<td>0–14</td>
<td>4</td>
<td>56</td>
<td>0–13</td>
<td>3</td>
<td>56</td>
<td>0–11</td>
</tr>
<tr>
<td>Nines rule</td>
<td>0</td>
<td>6</td>
<td>0–2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Unknown</td>
<td>1</td>
<td>11</td>
<td>0–6</td>
<td>0</td>
<td>6</td>
<td>0–2</td>
<td>0</td>
<td>6</td>
<td>0–2</td>
</tr>
</tbody>
</table>

*Note.* For mean use the percentage of problems on which each participant used each strategy was calculated and then averaged across the sample. Percentage of sample is the percentage of the entire sample who used a particular strategy at least once. Range of use is the highest and lowest percentage use of a particular strategy in the sample by a participant. CRT-R = choice reaction time task, random; SRT-R = simple reaction time task, random.
Table 2. Subtraction problem-solving response times and accuracies across working memory load conditions

<table>
<thead>
<tr>
<th>Strategy and problem size</th>
<th>Control</th>
<th>SRT-R (ms)</th>
<th>CRT-R (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval small</td>
<td>1,150 (68)</td>
<td>1,159 (78)</td>
<td>1,335 (81)</td>
</tr>
<tr>
<td>Retrieval large</td>
<td>1,393 (59)</td>
<td>1,389 (67)</td>
<td>1,621 (71)</td>
</tr>
<tr>
<td>Counting small</td>
<td>2,248 (175)</td>
<td>2,652 (256)</td>
<td>3,209 (441)</td>
</tr>
<tr>
<td>Counting large</td>
<td>3,746 (152)</td>
<td>4,492 (222)</td>
<td>6,195 (382)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strategy and problem size</th>
<th>Control</th>
<th>SRT-R</th>
<th>CRT-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval small</td>
<td>99 (1.6)</td>
<td>99 (1.7)</td>
<td>99 (2.3)</td>
</tr>
<tr>
<td>Retrieval large</td>
<td>94 (1.4)</td>
<td>94 (1.5)</td>
<td>91 (2.0)</td>
</tr>
<tr>
<td>Counting small</td>
<td>98 (1.3)</td>
<td>98 (1.7)</td>
<td>99 (2.2)</td>
</tr>
<tr>
<td>Counting large</td>
<td>98 (1.1)</td>
<td>95 (1.5)</td>
<td>90 (1.9)</td>
</tr>
</tbody>
</table>

Note. The numbers in parentheses are standard errors. Control = subtraction alone; CRT-R = choice reaction time task, random (dual task); SRT-R = simple reaction time task, random (dual task).

In the SRT-R condition, 1,211 ms longer, \( F(1, 55) = 17.88, p < .01 \), and those in the SRT-R condition were significantly longer than in the control condition, 599 ms longer, \( F(1, 55) = 11.13, p < .01 \). These ANOVAs were also conducted on the error data within each strategy separately, and no significant differences emerged, \( F_s < 2.75, ps > .05 \).

The next set of analyses was conducted to determine whether the counting strategy used more central executive resources than retrieval. In previous mental arithmetic experiments, problem size has been added as a variable in analyses, and these analyses have shown that nonretrieval strategy use, RTs, and errors all increase with the size of a problem. Problem size was included in analyses here to determine whether WM involvement would interact with strategy use or problem size. In accordance with the procedure of Seyler et al. (2003), problems with minuends of 11 or more were coded as large problems, and those with minuends less than 11 were coded as small problems. A 2 (problem size: small, large) × 2 (strategy: retrieval, counting) × 3 (load: control, SRT-R, CRT-R) ANOVA with repeated measures on the last two variables was conducted on both the subtraction RT and error item data. Table 2 shows the RTs and errors by strategy, problem size, and load condition. The ANOVA revealed significant main effects of problem size, \( F(1, 54) = 55.31, p < .001 \); strategy, \( F(1, 54) = 279.9, p < .001 \); and WM load, \( F(2, 108) = 28.65, p < .001 \). These main effects demonstrate that small problems were solved 1,180 ms faster than large problems, problems solved by retrieval were solved 2,415 ms faster than
problems solved by counting, and problems solved in the no load condition were solved 667 ms faster than problems in the SRT-R condition and 956 ms faster than problems in the CRT-R condition.

The aforementioned main effects were qualified by the problem size × strategy, $F(1, 54) = 41.26, p < .001$; problem size × load, $F(2, 108) = 4.83, p < .05$; strategy × load, $F(2, 108) = 17.69, p < .001$; and problem size × strategy × load, $F(2, 108) = 4.39, p < .05$, interactions. The latter two interactions were of particular interest. The strategy × load interaction resulted from the larger subtraction RT increase from the no load to both the SRT-R and CRT-R conditions for counting trials compared with retrieval trials (599-ms and 1,603-ms larger increase, respectively). For the three-way interaction, the difference in subtraction RTs across load conditions for large problems compared with small problems was greater for the counting strategy than for the retrieval strategy. Specifically, this effect resulted in the huge difference in RTs from small to large problems in the CRT-R condition for the counting strategy, $F(1, 54) = 4.14, p < .05$ (see Table 2). The same 2 (problem size: small, large) × 2 (strategy: retrieval, counting) × 3 (load: control, SRT-R, CRT-R) ANOVA was conducted on the error rate data. The only significant effect was the load × problem size interaction, $F(2, 108) = 3.12, p < .05$, indicating that, collapsing across strategies, the difference in accuracy between large and small problems was greater in the CRT-R condition.

Additional analyses were conducted to determine whether the different central executive loads interfered more with subtraction retrieval than number naming. A two-factor ANOVA was run with the within-participant factors of task (number naming, subtraction retrieval) and load (control, SRT-R, CRT-R). Significant main effects of task, $F(1, 378) = 704.89, p < .001$, and load, $F(2, 378) = 26.44, p < .001$, were found, but the task × load interaction was not significant, $F(2, 378) = 1.86, p > .15$, signifying that subtraction RTs did not differ more than number naming RTs across load conditions.

**SRT-R and CRT-R analyses**

Separate analyses of SRT-R and CRT-R data were conducted to compare secondary task performances alone with the concurrent number naming and subtraction problem-solving conditions. It should be noted that some secondary task condition data for two participants were lost through experimenter error, resulting in fewer degrees of freedom in the following analyses. Average tapping RTs for SRT-R alone, number naming with SRT-R, and subtraction with SRT-R were 130, 379, and 465 ms, respectively, and error rates (combined anticipatory and omission errors) were 3%, 23%, and 40%, respectively. A one-factor (load condition) within-participant ANOVA with RT as the dependent measure revealed a
significant effect of condition, $F(2, 32) = 123.45, p < .001$, and follow-up tests showed that all pairwise comparisons were significantly different, $Fs > 12.24, ps < .005$. The same one-way ANOVA conducted on the error measure yielded a significant effect of condition, $F(v2, 32) = 48.91, p < .001$, and follow-up tests showed that all pairwise comparisons were again significant, $Fs > 19.70, ps < .001$.

Turning to the CRT-R related tasks, average tapping RTs for CRT-R alone, number naming with CRT-R, and subtraction with CRT-R were 300, 444, and 471 ms, respectively, and error rates (combined anticipatory, omission, and decision errors) were 12%, 41%, and 62%, respectively. A one-factor (dual task condition) within-participant ANOVA with RT as the dependent measure revealed a significant effect of condition, $F(2, 32) = 26.02, p < .001$, and follow-up tests showed that all pairwise comparisons were significant, $Fs > 28.96, ps < .001$, with the exception of the CRT-R with number naming versus CRT-R with subtraction comparison, $F(1, 16) = 2.09, p > .15$. The same one-way ANOVA conducted on the error rate measure yielded a significant effect of condition, $F(1, 32) = 65.79, p < .001$, and follow-up tests showed that all pairwise comparisons were significant, $Fs > 39.10, ps < .001$.

**DISCUSSION**

The goal of this experiment was to further our understanding of the relationship between strategy use in mental arithmetic and involvement of the response selection subcomponent of the central executive of WM. More specifically, we predicted that a combined input monitoring and response selection task (CRT-R) would generate more interference than an input monitoring task (SRT-R) in both non–retrieval-based and retrieval-based problem solving, but the interference would be greater for non–retrieval-based problem solving.

Subtraction retrieval with the CRT-R dual task led to longer RTs than subtraction retrieval alone, the difference in RTs was greater for the CRT-R than the SRT-R dual task condition, and these effects were consistent across problem size. Neither the SRT-R nor the CRT-R task interfered more with number naming than subtraction retrieval RTs. Recall that the CRT-R task entails response selection (which button to press) and input monitoring (the tones were presented at random time intervals), whereas SRT-R entails only input monitoring. Thus, the findings indicate that response selection is involved in retrieval-based subtraction problem solving. The fact that the interference of the two load tasks was similar for subtraction retrieval and number naming may indicate that a major source of the interference from the CRT-R tasks (which was significant for both retrieval and number naming) may stem from
the initial attention associated with encoding the numeric stimuli. This possibility must be studied more carefully in future arithmetic and WM investigations.

As outlined previously, the current theoretical position on arithmetic fact retrieval is that when the two numbers of a problem are presented, nodes in an associative network in long-term memory are activated. In turn, activation spreads from these nodes to candidate answers, both correct answer nodes and incorrect near neighbor nodes (Galfano et al., 2003; Galfano, Mazza, Angrilli, & Umiltà, 2004; Rusconi et al., 2004). This spread of activation occurs automatically and appears to be unaffected by the addition of a central executive load (Rusconi et al., 2004). A second phase involves the selection of the answer from the set of candidate answers that have been activated. Because it involves a selection component, it is likely that this is the phase of arithmetic answer retrieval that would be disrupted by a secondary response selection task such as CRT-R. If that is the case, it appears likely that this interference is roughly equal across problems. This is supported by the lack of a problem size × load interaction that was found here and in other investigations (e.g., Deschuyteneer and Vandierendonck, 2005a, 2005b).

It would be interesting to study more systematically the possibility that the answer selection phase is where CRT-R interference occurs and to determine whether individual differences affect the degree and nature of the interference. For example, would the same pattern of results (in particular the lack of a problem size × load interaction) still occur if the associative strength of problem–answer connections were experimentally manipulated by having participants practice complex arithmetic problems to different levels of automaticity? One could imagine that people who had just started to retrieve answers to a set of arithmetic problems would have associations between problems and correct answers that are weaker, that are more variable across problems, and that are closer in magnitude to the strength of associations between problems and incorrect answers. In this example, the answer selection phase of retrieval would be more difficult and time consuming for these people because activation levels of correct and incorrect candidate answers should be poorly differentiated. As a result, a response selection central executive load might have a greater impact on problems with lower associative strengths than on those with higher associative strengths, resulting in a problem size × load interaction.

Although retrieval-based problem solving entails response selection related central executive resources, it is also clear that nonretrieval processes use these resources and use more than retrieval processes do. This was documented when counting-based strategies were used. RTs for counting trials were shortest for no load conditions, intermediate for SRT-R,
and longest for CRT-R dual task conditions, and these differences were larger than the differences for retrieval trials. Most interesting was the finding of a significant problem size × load interaction for counting trials: A CRT-R load increased counting RTs more for large problems than for small problems. One possible source of this interaction is that the CRT-R task interferes with processes central to executing a counting strategy; for example, counting rate may decrease (as was shown for addition problems by Hecht, 2002), or the ability to decide when the end of the count has been reached may be affected. Other sources of the interaction could be processes related to strategy selection that may be more resource demanding for larger problems. These possibilities include deciding whether counting up from the subtrahend or down from the minuend is most adaptive (efficient) and deciding whether counting is the most adaptive strategy to use for the current problem.

These possibilities are in accord with Siegler’s adaptive strategy choice model (Lemaire & Siegler, 1995) and more recent strategy choice and development simulation (Shrager & Siegler, 1998). During arithmetic skill development both answers and strategy use are associated with problems, and the strength of a particular strategy builds as information about its global success (speed and accuracy on all problems), success on subclasses of problems (e.g., problems with 9 as a subtrahend), and success on individual problems (e.g., 13 – 6) builds. Therefore, it would be beneficial in the future to compare central executive involvement in free choice and no-choice strategy conditions (where participants are forced to use one strategy on an entire set of problems). This would begin to allow researchers to determine central executive involvement during the strategy selection and strategy execution phases of arithmetic problem solving. The no-choice method also would allow secondary task performance to be compared across strategies to examine whether secondary task performance differences occur instead of, or in addition to, primary (arithmetic) task differences. In the present investigation most participants used a mix of retrieval and nonretrieval strategies, so secondary task performance could not be examined for each strategy separately.

Finally, it is important to test the involvement of other proposed subcomponents of the central executive (e.g., memory updating and inhibition) in arithmetic. The involvement of memory updating in arithmetic has been studied recently, with results indicating that addition and multiplication problem solving involve both response selection and memory updating processes (Deschuyteneer, Vandierendonck, & Muylleart, 2006). One drawback to the aforementioned study is that it did not involve the separation of trials into retrieval and nonretrieval solution procedures. Given the findings and arguments of Hecht (2002), Seyler et al. (2003), and the present investigation, it is imperative that researchers include
strategy use as a variable in WM investigations to get a more accurate and detailed picture of arithmetic processing.

Notes
Correspondence about this article should be addressed to Loel Tronsky, Psychology Department, 203 Weldon Hall, Albertus Magnus College, 700 Prospect Street, New Haven, CT 06511. Received for publication June 26, 2006; revision received December 30, 2006.

References


What we know and what we need to know. In J. M. Royer (Ed.), *Advances in mathematical cognition* (pp. 117–145). Greenwich, CT: Information Age Publishing.
