Counting on working memory in simple arithmetic when counting is used for problem solving

STEVEN A. HECHT
Florida Atlantic University, Davie, Florida

A concurrent-task methodology was used to investigate relations between the availability of aspects of working memory resources and both strategy selection and execution while simple addition equations (e.g., 4 + 3 = 8) were being verified. Consistent with prior research in which production trials have been used, undergraduates selected a variety of procedures other than retrieval. Availability of working memory resources did not generally affect strategy selection. Disrupting central executive and phonological aspects of memory affected strategy execution, but only when min counting was used to solve the problems. These and other features of the results suggest that availability of working memory resources does not contribute to individual differences in strategy selection and time to execute retrieval processes.

This study furthers our understanding of the potential contributions of working memory resources to strategy selection and computational processes during the solving of simple addition problems (e.g., 4 + 3 =) in adults. Working memory resources are required to solve multistep problems such as 125 + 97 (Fürst & Hitch, 2000; Logie, Gilhooly, & Wynn, 1994). In these problems, working memory is required for retaining intermediate values during calculation and for carrying across columns. Meanwhile, current process models do not make explicit claims about the influences of working memory on simple arithmetic problem solving (Hecht, 1999; Hecht, Torgesen, Wagner, & Rashotte, 2001). The goal of the present research was to address this limitation by exploring the role of working memory in simple arithmetic.

Working memory resources may influence both the selection and execution of strategies in simple arithmetic. Selection of strategies refers to the processes that occur before a particular procedure is executed (Bisanz & LeFevre, 1990). Most models predict that adults with extensive experience in solving arithmetic problems rely largely if not exclusively on direct retrieval (Ashcraft, Donley, Halas, & Vakali, 1992; Groen & Parkman, 1972; Kirk & Ashcraft, 2001; Siegler & Shipley, 1995). Contrary to that claim, many adults use nonretrieval strategies (Geary & Wiley, 1991; Hecht, 1999; LeFevre, Sadesky, & Bisanz, 1996). For example, Hecht asked participants to say the answers to simple addition problems (i.e., production trials). He found that over half of the adult sample (*n* = 61) used a counting strategy at least once. Thus,

current process models need to account for multiple strategy usage in adults (Hecht, 1999). The current mechanisms assumed by Siegler and Shipley's (1995) adaptive strategy choice model (ASCM; see also Shrager & Siegler's, 1998, SCADS model) may explain multiple strategy usage in adult simple arithmetic performance. The ASCM assumes that a particular strategy tends to be selected when that procedure can be executed efficiently (i.e., with sufficient speed and efficacy). Thus, even when memory resources are substantially unavailable, an adult might use a nonretrieval strategy for a problem if that procedure can be executed efficiently (Hecht, 1999).

Execution of strategies refers to processes that occur when a particular strategy is used to solve a simple arithmetic problem. Current process models do not explicitly posit availability of memory resources as a direct influence on strategy execution. It is likely that the existing mechanisms of current process models are sufficient for explaining retrieval processes. Retrieval appears to generally occur automatically, without requiring substantial memory resources (see, e.g., Craik, Govoni, Naveh-Benjamin, & Anderson, 1996; LeFevre, Bisanz, & Mrkonjic, 1988). Meanwhile, counting-based procedures in simple arithmetic appear to require substantial memory resources in children (Adams & Hitch, 1997). Counting requires memory resources for monitoring the number just said and the next count (Case, 1985; Logie & Baddeley, 1987).

Ashcraft et al. (1992) proposed using concurrent task methodology to study the possible role of working memory in performing mental arithmetic. In their study, the same participants solved problems via the verification task. In verification, participants verify the truth of an equation (e.g., 3 + 4 = 8 should be indicated as false). The articulatory suppression condition involved verification while a letter was repeated at a "rapid and constant rate." Articulatory suppression consumes memory resources devoted to the phonological loop (Baddeley,

1992). Two other conditions consumed resources devoted to the central executive. The central executive is associated with a variety of activities such as performance monitoring and coordination of activities (Baddeley, 1996; Engle, Tuholski, Laughlin, & Conway, 1999).

Ashcraft et al. (1992) found that mean raw verification times were significantly longer (by about 400 msec) in both central executive load conditions than in the articulatory suppression condition. Interpretation of participants' mean time to verify simple arithmetic problems (i.e., raw verification times) can be misleading. Increases in raw verification times may reflect increased time to execute strategies or other processes involved in verification. As pointed out by Campbell and Tarling (1996), raw verification times reflect processes involved in encoding the problem addends and answer, computing the answer, comparing the computed answer with the stated answer, and decision processes regarding which button to press for indicating whether the stated answer is true or false. Interpretations from raw verification times are also tenuous when no differences in raw latencies emerge between conditions. If no differences are found, it may indicate that computation processes were unaffected. It is also possible that some processes involved in verification (e.g., decision processes associated with pressing a button) were sped up as a result of additional time spent on computation processes. Kaye, deWinstanley, Chen, and Bonnefil (1989) reported evidence suggesting that raw latencies do not reflect how availability of memory resources contributes to verification during the solving of two-term additions (e.g., 34 + 56). Lack of available memory resources affected the computational phase in children and affected processes associated with comparison and decision stages in adults.

Ashcraft et al. (1992) disentangled the effects of concurrent load on strategy execution processes and other component processes involved in verification. For each condition, they used regression procedures to estimate the time for retrieval processes to be implemented. They did not report results for false equations. Next, they determined whether retrieval processes took longer to be implemented between conditions. Their analyses revealed no significant differences between articulatory suppression and the other two conditions assumed to require both phonological loop and central executive resources. Another study, based on two experiments, reported no significant effects of concurrent memory load on estimated time to execute retrieval processes (Lemaire, Abdi, & Fayol, 1996). Although only 5 participants were included in each concurrent load condition, Lemaire et al.'s results do provide converging evidence against the claim that substantial memory resources are used while simple arithmetic is being solved.

In the present study, undergraduates verified simple addition equations in three within-participants conditions. A control condition was included in which the participants solved problems silently. In the articulatory suppression condition, the participants repeated the same letter during each verification trial, which consumes memory resources devoted to the phonological loop. Finally, the participants verified equations while they randomly generated sequences of letters, which consumes memory resources devoted to both the phonological loop and the central executive (Baddeley, 1996; Teasdale et al., 1995).

The primary focus of the present study was to answer two questions regarding the role of working memory in simple addition. The first question was whether availability of working memory resources influences the selection of strategies in simple arithmetic. The strategies that the adults used to solve the problems were determined on a trial-by-trial basis. This made it possible to determine whether secondary tasks influenced strategy selection. This study also examined whether individual differences in working memory ability were associated with strategy selection. By recording strategies, this study also sheds light on another important issue in the math cognition area—that is, do adults use similar strategies to solve verification and production trials? Some have argued that verification processes are different from the strategies used in production trials (e.g., Campbell & Tarling, 1996; Krueger & Hallford, 1984; Lemaire & Fayol, 1995). In particular, participants might rely on plausibility judgments to determine whether the problem and stated answer are correct, rather than actually determining the answer to the problem.

The second question was whether availability of memory resources influences the execution of simple arithmetic strategies. This possibility was investigated by using improved methods in two important ways. Findings regarding solution times averaged across strategies can be considerably misleading if execution of only some strategies requires substantial memory resources (cf. Siegler, 1987). All previous reports have failed to segregate trials by strategy. Thus, the effects of working memory load on execution of strategies might have been underestimated. The second methodological improvement was that both articulatory suppression and random letter generation conditions were implemented in a more controlled manner than in previous studies. In the Ashcraft et al. (1992) study, participants were simply told to say the letters "at a rapid and constant rate." In the Lemaire et al. (1996) study, participants were instructed to say a letter every 2 sec. Before the experimental trials began, participants practiced saying a letter every 2 sec with the aid of a metronome beat. In the present study, measures were taken to ensure that the articulatory suppression and random letter generation conditions were faithfully implemented by the participants. This study also determined whether the pattern of results concerning strategy selection and execution depended on how well the participants followed the secondary task instructions.

METHOD

Participants

Thirty-four undergraduate psychology students at Florida Atlantic University participated in this experiment for extra course credit.

Materials

Simple arithmetic–silent control condition. The participants verified simple addition equations silently. All stimuli were presented in standard horizontal form ($a + b = c$). None of the problems contained an addend of zero or identical addends (i.e., ties). The *true* stimuli were 36 equations with correctly stated answers, drawn from one of two stimulus sets. Approximately half of the participants were given Stimulus Set A. Both sets included one of two possible combinations of single digit integers from $1 + 2$ to $8 + 9$. The operand order for any given problem in Stimulus Set A (e.g., $4 + 5$) was the opposite of that given in Stimulus Set B (i.e., $5 + 4$). For each stimulus set, half of the items were presented with the minimum operand left of the operation sign. Presentation order of stimuli was randomized independently for each participant within each condition. The *false* stimuli were 36 equations with incorrectly stated answers, drawn from Stimulus Set A or B. The participants who were administered Stimulus Set A for true problems were given Stimulus Set B for false problems, or vice versa. Incorrectly stated answers were chosen so that the difference between the correct and stated answer (i.e., split) would not influence the pattern of results. Splits were plus or minus one or two. The correlation between split and problem solution times, averaged across participants approached zero ($r = .06$).

Simple arithmetic–articulatory suppression condition. This subtest was identical to the control condition, except that problems were verified while the participants were repeating orally the same letter of the alphabet.

Simple arithmetic–random letter generation condition. This subtest was identical to the control condition, except that problems were verified while the participants were saying different letters in random fashion.

Working memory. This was the operation span with words task (Engle et al., 1999). The participants were presented individual alphabet–arithmetic–word strings (e.g., $d + 2 = h$ milk). On each trial, the participants indicated whether the stated answer was true or false. For example, $d + 2 = h$ is false because f is the answer (i.e., d, e, f). Immediately after stating true or false, the participants read aloud the given word. On hearing the stated word from the participant, the experimenter pressed a key, presenting the next string. After the last alphabet–arithmetic–word string in the set, the participants stated each of the words in that set in the order that they were presented. Engle et al. (1999) asserted that tasks like operation span with words substantially capture memory processes devoted to both the phonological loop and the central executive. The number of alphabet–arithmetic–word strings (set size) varied from two to six. There were 2 practice trials and 10 test trials. All students were over 85% accurate on the alphabet–arithmetic portion of the test. Partial credit was given in order to gain maximum variability on this task (Torgesen & Houck, 1980). That is, each word recalled in the correct order was given one point, and each word recalled in the incorrect order was assigned half a point. The maximum score that could be obtained was 40. Mean total points was 21.15 ($SD = 4.85$).

General computation. This task was the calculation subtest from the Woodcock–Johnson Psycho-Educational Battery (Woodcock & Johnson, 1990), which provides an estimate of the total number of computation algorithms mastered. The participants were told to answer the problems as accurately as possible, to skip items that they could not answer, and to take as much time as needed. There were 58 items, and mean correct score was 38.32 ($SD = 4.58$).

Procedure

Each participant was individually tested, and the order of all subtests was randomized independently for each. The participants were permitted a short rest period between subtests. The silent control, articulatory suppression, and random letter generation subtests each involved a total of 5 practice trials and 72 test trials. The general procedure for each verification problem was as follows: Each trial began with a series of six characters presented horizontally at the middle of the screen for 2,000 msec. Next, the problem and stated answer were presented in standard horizontal form (e.g., $1 + 2 = 4$) in the center of the monitor. The equations stayed on the screen until the participant pushed either the right shift or left shift key. About half of the participants pushed the left shift key if the problem and stated answer were true and the right shift key if they were false. The other half pushed the right shift key for true and the left shift key for false responses. The participants were told to push the appropriate button as quickly and accurately as possible. The time interval between when the problem was presented and when the button was pressed was recorded by the computer with millisecond accuracy. Immediately after the button was pressed, the experimenter asked the participant how he or she solved the problem and wrote verbatim the verbal description. Except for one sentence, the wording of the present instructions for reporting solution methods was taken from Hecht (1999). The participants were told that retrieval, counting, or a "special trick" could be used to solve addition problems. A new sentence stated, "It is even possible that we may just know that the answer is correct or incorrect without even having to solve the problem—the answer just doesn't look right."

In the control condition, the trial began with a series of pound signs (i.e., # # # # #). Next, the participants were presented the arithmetic problem and the stated answer. The participants pushed the appropriate key silently. For the articulatory suppression condition, the trial began with presentation of a series of six identical capital letters (e.g., A A A A A A). The letter to be repeated was either A, B, C, D, E, or F. Each arithmetic problem and stated answer was randomly assigned a particular letter. The participants were instructed to repeat the presented letter aloud as soon as the letter string was presented each time they heard a click. The participants heard a click every 895.5 msec until they pushed a button. To repeat letters at a rate of approximately once per second is considered to demand considerable working memory resources (Marsh & Hicks, 1998; Teasdale et al., 1995). For the random letter generation condition, the procedure was identical to the articulatory suppression trials, except for the following difference. The trial began with presentation of Series A–F (i.e., A, B, C, D, E, F). The participants were instructed to say without stopping a random series of letters from the series A–F in tandem with the metronome click. They were instructed to say a different random sequence for each trial. The uttered letters were recorded via a tape recorder so that the participants' adherence to the requirement to say a letter at each click could be verified.

Coding of Solution Procedures and Adherence to Concurrent Memory Load Tasks

The protocols were coded on the basis of the categories used by LeFevre, Sadesky, and Bisanz (1996). Responses were coded as *retrieval* when the participants claimed that they remembered the answer. *Decomposition* procedures involved the use of a known arithmetic fact to find an answer. For example, a participant might solve the problem "$6 + 7$" by finding the solution to "$6 + 6$" and then adding 1 to 12. *Min counting* involved counting from the larger addend the number of increments of the smaller addend (e.g., "5 plus 2 equals 5, 6, 7; 7 is the answer"). Finally, an *other* category was used to code responses that were not consistent with the other substantive strategies (e.g., memory aids such as visualizing the answer in a fact table, guessing, hand trick with fingers, etc.). Also included was a new category called *implausibility judgement*, in which it was reported that the equation "just looked incorrect" (Lemaire & Fayol, 1995).

If the participants adhered to the concurrent task instructions, the total number of letters said per trial should approach the number of times the participants were expected to say a letter (i.e., every 895.5 msec). For each trial, the expected number of letters said was equal to trial duration divided by 895.5. Following Hegarty, Shah, and Miyake (2000), percent adherence (i.e., ratio of actual to expected) is the rate of producing letters so that any values smaller than 100% would indicate the degree of failure to faithfully produce the letters in tandem with the metronome.

RESULTS

Working Memory and Selection of Strategies

The first question addressed by this study was whether availability of memory resources influenced the selection of strategies in simple arithmetic. The percentage of trials in which each participant used each strategy was determined. Mean percentages of strategy selection averaged across participants, segregated by condition and truth status, are shown in Table 1. Analyses were conducted to determine whether the participants' strategy selection differed between conditions. For true trials, no significant differences emerged in the frequency with which the participants used each substantive strategy between conditions. For false trials, significant differences in percentage of trials in which the participants used the implausibility strategy emerged between the silent, articulatory suppression, and random letter generation conditions [9%, 13%, and 15%, respectively; $F(2,32) = 4.645, p < .02$]. The participants used the implausibility judgment strategy more often in the articulatory suppression [$F(1,33) = 6.379, p < .02$] and random letter generation conditions [$F(1,33) = 6.161, p < .01$] than in the control trials.

Strategy choices may have been impacted by concurrent memory load, but only for trials in which the participants adhered most to the secondary task demands. Strategy choices were examined for both high and not-so-high percent adherence trials, on the basis of the median split of secondary task adherence. Compared with control trials, comparable frequencies of strategy selection emerged for both high and not-so-high adherence trials. Strategy choices may have been associated with individual differences in working memory ability. Correlations between memory skills and variability in arithmetic performance are shown in Table 2. No relations emerged between working memory and variability in strategy selection. Interestingly, working memory did correlate with general math computation ($r = .52, p < .001$), suggesting that higher order math is influenced by memory ability. Engle et al. (1999) reported a similar correlation between adults' operation span with words and quantitative SAT ($r = .46$; see also Ashcraft & Kirk, 2001; Geary & Widaman, 1992). In sum, converging evidence was found that strategy selection was not influenced by availability of memory resources.

Working Memory and Execution of Strategies

The second question addressed by this study was whether the availability of memory resources influenced the execution of arithmetic strategies. There were four dependent measures of strategy execution: accuracy, raw latencies, degree of adherence to secondary tasks, and estimated time to engage in computational processes. In the analyses pertaining to a particular strategy, mean performance for each equation was based on trials in which the participants reported using that strategy. Mean performance was obtained for each equation by averaging across participants (i.e., item analysis) instead of con-

sidering performance of the individual participants (i.e., participants analysis). Item analyses were conducted because predictors, assumed to capture computational processes, provide a much better account of averaged data than does the variability in individual participant's latencies (Hecht, 1999). Note that slope values obtained in both item and participants analyses are comparable in magnitude (Lorch & Meyers, 1990). Item analysis also enables comparisons of performance between conditions to be, in general, based on similar numbers of items per condition. Item analysis is typically used in investigations of mental processes involved in simple arithmetic (see, e.g., Ashcraft et al., 1992; Geary & Wiley, 1991; Groen & Parkman, 1972; Hecht, 1999; LeFevre, Sadesky, & Bisanz, 1996; Siegler & Shipley, 1995).

Accuracy

Although adults rarely make errors in simple arithmetic, accuracy reflects the effectiveness of executing computational processes while a particular strategy is being used (Campbell & Timm, 2000). Table 1 depicts mean accuracy, segregated by condition and truth status. For both true and false trials, there were significant differences ($ps < .05$) in the overall accuracy between conditions [$F(2,105) = 3.195$ and $F(2,105) = 3.777$]. For true trials, overall accuracy was significantly ($ps < .05$) reduced in comparison with control performance for both articulatory suppression [$F(1,70) = 5.460$] and random letter generation [$F(1,70) = 5.131, p < .05$]. For false trials, overall accuracy was significantly reduced in comparison with control performance for both articulatory suppression [$F(1,70) = 5.374, p < .05$] and random letter generation [$F(1,70) = 8.284, p < .01$]. Accuracy differences did not emerge between the articulatory suppression and random letter generation conditions. Accuracy differences also did not emerge when accuracy was considered separately for trials in which each strategy was used. In sum, these analyses provide some support for the idea that reduced availability of attentional resources impairs the efficiency with which adults execute simple arithmetic strategies.

Raw solution times. Recall that raw latencies reflect time to execute all component processes involved in verification (Campbell & Tarling, 1996). Mean raw latencies are reported in Table 1. For both true and false retrieval trials, significant differences ($ps < .01$) emerged between the control and random letter generation conditions [$F(1,66) = 48.132$ and $F(1,70) = 24.748$, respectively]. For counting trials, significant differences emerged between the control and random letter generation conditions for true trials [$F(1,65) = 9.479, p < .01$] but not for false trials [$F(1,68) = 0.288, p > .10$]. For decomposition trials, significant differences emerged between the control and random letter generation conditions for true trials [$F(1,35) = 4.693, p < .05$] but not for false trials [$F(1,33) = 0.481$]. For both true and false other trials, significant differences ($ps < .05$) emerged between the control and random letter generation conditions [$F(1,34) = 5.608$ and $F(1,49) = 5.570$, respectively]. No significant differ-

Table 1
Selection of Procedures, Latencies, and Variability of Latencies, Segregated by Strategy, Condition, and Truth Status

Procedure	Mean % Use		% of Sample	Mean % Correct		Range of % Use	Mean RT		Mean % Adherence	
	M	*SD*		*M*	*SD*		*M*	*SD*	*M*	*SD*
True Verification Trials: Silent Control										
Retrieval	70	23	100	97	4	6–100	1,162	141	–	–
Decomposition	11	10	71	96	11	0–29	2,385	410	–	–
Min counting	14	20	56	98	5	0–75	1,852	406	–	–
Other	2	5	29	82	32	0–29	1,881	670	–	–
Overall errors	4	4	71	–	–	0–15	1,814	988	–	–
True Verification Trials: Articulatory Suppression										
Retrieval	69	22	97	95	5	33–100	1,151	163	98	4.6
Decomposition	11	10	71	95	10	0–39	2,257	592	88	8.1
Min counting	12	17	56	95	10	0–67	1,827	829	85	16
Other	4	6	41	66	42	0–25	1,980	898	91	19
Overall errors	7	5	91	–	–	0–19	2,055	1,470	85	15
True Verification Trials: Random Letter Generation										
Retrieval	69	24	100	97	6	11–100	1,466	183	90	6.3
Decomposition	12	11	71	85	29	0–33	2,741	570	81	9.8
Min counting	12	20	56	95	11	0–75	2,332	802	77	12
Other	4	5	47	82	34	0–22	2,601	1,029	73	24
Overall errors	6	6	67	–	–	0–25	2,218	1,353	81	21
False Verification Trials: Silent Control										
Retrieval	55	29	100	94	6	0–97	1,492	201	–	–
Decomposition	8	10	71	99	4	0–36	2,732	230	–	–
Min counting	13	19	68	98	6	0–72	2,396	902	–	–
Implausibility	9	12	77	100	1	0–39	1,556	544	–	–
Other	9	17	85	79	34	0–92	1,507	598	–	–
Overall errors	5	5	77	–	–	0–17	1,733	959	–	–
False Verification Trials: Articulatory Suppression										
Retrieval	54	26	100	95	6	3–100	1,402	193	94	6.3
Decomposition	8	9	59	90	25	0–31	2,778	233	82	9.6
Min counting	10	15	53	90	24	0–64	2,366	1,096	82	17
Implausibility	13	15	62	98	5	0–47	1,601	530	91	10
Other	9	15	68	81	28	0–83	1,397	527	98	50
Overall errors	6	5	79	–	–	0–17	1,819	1,084	90	13
False Verification Trials: Random Letter Generation										
Retrieval	50	30	91	89	18	0–100	1,639	190	88	13
Decomposition	7	8	56	88	17	0–23	2,911	253	76	20
Min counting	11	17	47	91	22	0–81	2,494	865	76	10
Implausibility	15	20	44	97	7	0–69	2,024	624	88	10
Other	8	14	47	87	20	0–83	1,874	1,094	85	13
Overall errors	8	7	79	–	–	0–28	1,991	1,190	77	15

Note–Mean % Use, average percentage of trials in which the participants used that strategy; % of Sample, percentage of sample that used that strategy at least once or made an error at least once; Mean % Correct, percentage of trials in which participants correctly verified the answer; Range of % Use, range of percentage of trials in which individual participants used that strategy; Mean RT, average solution times for trials using that strategy, which are measured in milliseconds; Mean % Adherence, average percentage of letters that were said relative to the expected production of letters. Retrieval, trials in which participants claimed that they remembered the answer. Decomposition, the use of a known arithmetic fact to find an answer. For example, the student may solve the problem "6 + 7 = 13" by finding the solution to "6 + 6" and then adding 1 to 12. Min Counting, trials in which participants counted from the larger addend the number of increments of the smaller addend (e.g., "5 plus 2 equals 5, 6, 7; seven is the answer"). Other, any other strategies that were not consistent with the other substantive strategies and were executed in less than 1% of the trials (e.g., memory aids such as visualizing the answer in a math fact table, guessing, hand trick using fingers, etc.). Overall Errors, trials, across strategies, in which participants answered incorrectly.

ences in raw latencies were found between the control and articulatory suppression conditions. Although sparse, some previous reports have also suggested that raw control and articulatory suppression times do not differ (De Rammelaere, Stuyven, & Vandierendonck, 1999, 2001; but see Lemaire et al., 1996). In sum, these results suggest that a memory load on the central executive disrupted component verification processes.

Adherence to secondary tasks. Articulatory suppression and random letter generation require substan-

Table 2
Intercorrelations Between Working Memory, General Math Computation, and
Individual Patterns of Simple Arithmetic Performance (n = 34)

Working memory	–							
General computation	.52**	–						
Retrieval use	−.02	.07	–					
Counting use	.06	−.01	−.77†	–				
Decomposition use	−.17	−.01	−.48**	.02	–			
Other use	.21	.09	−.35*	−.01	.07	–		
Implausibility use	−.09	−.08	−.55†	.25	.20	−.05	–	
Estimated time for retrieval processes	.21	.17	.13	.05	−.43**	.15	−.23	–
Accuracy	−.02	.44**	.13	.07	−.21	.08	−.04	−.07	–

Note—Simple arithmetic performance was based on all correctly answered true and false trials except implausibility strategy use and time for retrieval processes. Implausibility use was based on false trials only. Estimated time for retrieval processes are individual participants' retrieval use slopes based on true retrieval trials only.   *$p < .05$.   **$p < .01$.   †$p < .001$.

tial working memory resources (Hegarty et al., 2000; Torgesen, 1996). Thus, performance on these tasks might be reduced when attentional resources are consumed by a memory-demanding arithmetic strategy (cf. McNamara & Scott, 2001). Mean percent secondary task adherence, segregated by strategy, is reported in Table 1. Significant differences ($ps < .001$) between strategies in percent adherence emerged for true articulatory suppression and for both true and false random letter generation trials [$F(4,155) = 5.604$, $F(4,160) = 4.907$, and $F(5,185) = 11.51$, respectively]. Percent adherence was consistently greater for retrieval than for both counting and decomposition trials ($ps < .01$ or $.05$). This suggests that retrieval involved the smallest tradeoff between secondary task adherence and strategy execution.

Computational processes. The time involved in execution of arithmetic strategies was examined. Following Ashcraft et al. (1992), regression procedures were used to examine relations between predictors and item variability in verification times. LeFevre, Bisanz, Daley, Buffone, and Sadesky (1996) proposed an associative strength predictor variable representing the distribution of associative strengths for a given problem (referred to as retrieval use). Retrieval use for each problem was calculated by dividing the number of participants who used retrieval for that problem by the total sample size (Hecht, 1999; LeFevre, Bisanz, et al., 1996). The min counting model assumes that the size of the smallest addend captures the time involved in each increment of counting needed to solve a problem (Groen & Parkman, 1972).

Table 3
Slope (S) in Milliseconds and R^2 Values, Segregated by Strategy, Condition, and Truth Status

	All		Retrieval		Decomposition		Counting		Implausibility Judgment	
Addition Strategy	S	R^2	S	R^2	S	R^2	S	R^2	S	R^2
True Verification Trials: Silent Control										
Retrieval use	−14.81	.77	−5.19	.46	17.50	.52	−13.66	.40	–	–
Min	126.17	.65	52.23	.54	−30.90	.02*	126.85	.37	–	–
True Verification Trials: Articulatory Suppression										
Retrieval use	−16.73	.70	−5.40	.33	20.20	.13*	−20.50	.35	–	–
Min	146.15	.71	58.0	.51	143.81	.16*	232.60	.62	–	–
True Verification Trials: Random Letter Generation										
Retrieval use	−19.67	.80	−6.27	.38	−6.106	.01*	−30.80	.43	–	–
Min	160.25	.66	62.79	.47	130.02	.11*	280.36	.36	–	–
False Verification Trials: Silent Control										
Retrieval use	−19.44	.47	−5.96	.19	−0.943	.04*	−23.39	.28	−13.75	.19
Min	174.52	.74	52.05	.22	37.07	.00*	182.14	.36	121.47	.26
False Verification Trials: Articulatory Suppression										
Retrieval use	−18.80	.53	−4.22	.13	36.05	.17	−43.47	.38	−18.50	.55
Min	178.51	.74	65.86	.49	16.49	.00*	413.97	.56	124.70	.35
False Verification Trials: Random Letter Generation										
Retrieval use	23.69	.60	−4.32	.11	−24.42	.08*	−25.01	.16	−19.13	.25
Min	180.06	.72	46.44	.24	115.61	.05*	289.43	.46	181.35	.46

Note—All R^2s significant ($p < .05$) unless otherwise indicated. Retrieval use stands for the proportion of participants in the sample that used retrieval to solve each simple arithmetic problem. Min stands for the minimum addend in each problem.   *not significant.

Slope estimates and percentage of variance explained, segregated by condition, strategy, and truth status are reported in Table 3. The slope represents the time (in milliseconds) required to execute computational processes captured by the predictor (Ashcraft et al., 1992; LeFevre, Bisanz, et al., 1996). The retrieval use and minimum addend slopes for decomposition trials were not interpreted because, consistent with Hecht (1999), the R^2 values were not consistently different from zero.

In order to compare the present results with those of Ashcraft et al. (1992), slope values were first obtained without segregating trials by strategies. For both true and false trials, slope values were only slightly greater in the articulatory suppression and random letter generation conditions than the silent control condition. The magnitude of differences in the present slopes for all true trials between the articulatory suppression and the random letter generation conditions (2.94) was very similar to that obtained by Ashcraft et al. for true trials. They reported differences in associative strength slopes of 2.23 and 2.32 between articulatory suppression and two concurrent conditions assumed to load both the phonological loop and central executive. The strongest test for effects of memory load on execution of retrieval processes was obtained by examining retrieval use slopes for only retrieval trials. For both true and false equations, there were no significant differences in retrieval use slopes between conditions (ps >.10). The present differences in slopes between conditions were at most less than 2 msec, suggesting that retrieval processes were executed without requiring substantial memory resources.

Possible effects of concurrent memory load on counting processes were examined. For all trials, the secondary tasks did not affect the magnitudes of minimum addend slopes. The strongest test for possible effects of concurrent memory load on counting processes was obtained by determining slope values based on counting trials only. For counting trials, differences in minimum addend slopes between the silent, articulatory suppression and random letter generation conditions approached significance for true trials and was significant for false trials [$F(2,94) = 3.006, p < .054$ and $F(2,94) = 4.931, p < .009$, respectively]. Planned comparisons for true and false trials indicated that counting processes were significantly slowed down in comparison with those in the control trials for articulatory suppression [$F(1,64) = 6.872, p < .01$ and $F(1,64) = 13.573, p < .001$, respectively]. Planned comparisons for true trials indicated a significant slow down in counting in the random letter generation condition compared with in the control condition [$F(1,62) = 4.580, p < .04$]. Compared with the false silent control condition, there was an additional estimated 107 msec of processing time needed for each increment of counting in the false random letter generation condition. However, this difference in slopes did not approach conventional levels of significance [$F(1,65) = 1.136, p > .10$]. For both true and false problems, counting processing time was not significantly greater in the random letter generation

condition than in the articulatory suppression condition (ps > .10). In general, these results provide strong evidence that counting processes require substantial working memory resources.

Next was determined whether slope magnitudes depended on how well the participants adhered to the secondary task demands. To interpret any significant interaction, the trials were separated into two groups on the basis of a median split of degree of task adherence, and slope values were reestimated for these two groups, separately. For both true and false problems, slope values based on all trials and only retrieval trials did not interact with percent adherence (ps > .10). For true counting trials, secondary task adherence interacted with estimated time to engage in counting processes (i.e., min slopes) for random letter generation trials [$F(1,30) = 13.183, p < .001$]. For true random letter generation counting trials, estimated slope values for the good and not-so-good secondary task adherence groups were 154.83 and 305.03, respectively. For false counting trials, min slopes interacted with secondary task adherence within both articulatory suppression and random letter generation conditions [$F(1,29) = 8.339, p < .01$, and $F(1,29) = 11.233, p < .001$, respectively]. For false articulatory suppression counting trials, estimated slope values for the good and not-so-good groups were 292.38 and 693.04, respectively. For false random letter generation counting trials, the median split revealed slope values for the good and not-so-good groups of 209.99 and 508.59, respectively. When degree of task adherence was considered, differences in slope values between the false silent control and the random letter generation trials emerged. That is, the estimated counting processing time of 508.59 msec was significantly greater than the slope value of 182.14 obtained in the control trials [$F(1,48) = 4.770, p < .05$]. In sum, counting processes were increasingly disrupted by memory load when the degree of secondary task adherence decreased. Interestingly, this suggests that when limited attentional resources were consumed during counting trials, performance on both primary and secondary tasks tended to be impaired (cf. Hegarty et al., 2000).

DISCUSSION

The first question addressed in this study was whether availability of working memory resources influences the selection of strategies in simple arithmetic. A clear pattern of comparable strategy selection was found between conditions. Degree of secondary task adherence did not affect the pattern of strategy choices. The correlational results were also consistent with the idea that availability of working memory resources does not influence strategy selection in simple arithmetic. Interestingly, frequencies of strategy use for true equations were comparable with those of previous reports in which production trials were used and provide additional support for the idea that adults use multiple strategies in simple arithmetic

(Geary & Wiley, 1991; Hecht, 1999; LeFevre, Bisanz, et al., 1996). Only false trials provided some support for the idea that adults use a qualitatively distinct strategy to verify some arithmetic problems when compared with production trials (cf. Campbell & Tarling; 1996; De Rammelaere et al., 1999; Lemaire & Fayol, 1995). That is, across conditions, adults used the implausibility judgment strategy about as frequently as they did min counting (i.e., 13% and 11% of the time, respectively).

What mechanism can account for multiple strategy usage in adult simple arithmetic? Hecht (1999) hypothesized that perhaps many adults continue to use counting-based strategies because these procedures can be implemented very efficiently (automatically), like retrieval. Indeed, the ASCM assumes that a particular strategy is likely to be selected if that procedure can be executed efficiently. This hypothesis was not supported, because the min counting strategy required substantial working memory resources in order to be executed, but it was selected just as often between conditions. Thus, simple arithmetic process models do not appear to need additional mechanisms for how availability of memory resources influences strategy selection in adults. Other explanations are clearly needed to explain multiple strategy usage in adults. For example, Kirk and Ashcraft (2001) demonstrated that instructions given to participants for obtaining verbal protocols of strategy usage can be purposely worded in a way that biases adults to report either more or less reliance on nonretrieval strategies. They argued that it is possible that even relatively neutral instructions, such as those used in the present and previous investigations, might somehow bias adults to report multiple strategies.

The second question in the present study was whether availability of limited attentional resources influences the execution of simple arithmetic strategies. When all trials were considered, the results appear to support the conclusion that working memory load does not affect computational processes. However, segregation of trials on the basis of strategy use made it possible to determine when the participants used substantial memory resources during on-line simple arithmetic problem solving (cf. Logie, Della Sala, Laiacona, Chalmers, & Wynn 1996; Siegler, 1987).

Three converging sources of evidence suggest that substantial memory resources are involved in executing the counting strategy. First, the most crucial evidence is that counting processes were significantly slowed down during the secondary task conditions. That is, slope values reflecting time to execute counting processes were substantially longer for both the articulatory suppression and the random letter generation conditions than in the silent control condition. Second, although secondary task adherence was overall very good, adherence was significantly poorer when the participants used the counting strategy than when they used the automatically implemented retrieval strategy. Thus, secondary task performance was disrupted when substantial working memory resources were used to implement the counting

strategy. Finally, there is evidence that the results were replicated between the true and false equations.

Four converging sources of evidence suggested that retrieval processes did not involve substantial attentional resources devoted to the phonological loop and central executive. First, compared with that in the control trials, estimated time to execute retrieval processes was not affected by either secondary task. Importantly, retrieval slope values did not interact with secondary task adherence. Thus, there was no increased disruption of automatic retrieval processes when the participants most faithfully allocated limited attentional resources toward performing the memory-demanding secondary tasks. Second, secondary task percent adherence was significantly greater for retrieval trials than for both min counting and decomposition trials. Presumably, secondary task performance was enhanced because there was negligible competition for working memory resources between automatic retrieval and attention-demanding secondary task processes (cf. Gilhooly, Logie, Wetherick, & Wynn, 1993; Hegarty et al., 2000). Third, the present null results regarding the relations between working memory ability and variability in estimated time to engage in retrieval processes provides evidence that retrieval occurs automatically and is not substantially affected by memory constraints. Finally, the pattern of results was replicated for both true and false problems. In sum, the present pattern of results adds to the emerging consensus, based on findings from a variety of tasks, that retrieval tends to be implemented automatically (Baddeley, 1992, 1996; Craik et al., 1996; LeFevre et al., 1988; Zbrodoff & Logan, 1986).

Evidence pointing to a general effect of random letter generation on raw verification latencies was obtained. Raw verification latencies are determined by strategy execution time and also by other components of performance, particularly comparison and decision processes (Campbell & Tarling, 1996). Interestingly, raw verification times, even for retrieval trials, were significantly slower in the random letter generation condition than in the control condition. It is not surprising that random letter generation affected the raw latencies, because comparison and decision processes are defining characteristics of the central executive aspect of working memory (Torgesen, 1996). Although sparse, previous work has also suggested a general effect of random letter generation on raw verification times when compared with a silent control task (see De Rammelaere et al., 1999, 2001; Lemaire et al., 1996).

The present results also have important implications for current models of multistep arithmetic. Articulatory suppression and random letter generation methodology reveals that phonological loop resources are used to store intermediate values, whereas carrying operations require central executive resources (Fürst & Hitch, 2000; Logie et al., 1994). The present results suggest that substantial memory resources are also used for computing each intermediate step, but only when participants use a strategy that involves counting.

REFERENCES

ADAMS, J. W., & HITCH, G. J. (1997). Working memory and children's mental addition. *Journal of Experimental Child Psychology*, **67**, 21-38.

ASHCRAFT, M. H., DONLEY, R. D., HALAS, M. A., & VAKALI, M. (1992). Working memory, automaticity, and problem difficulty. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 301-329). Amsterdam: Elsevier.

ASHCRAFT, M. H., & KIRK, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General*, **130**, 224-237.

BADDELEY, A. D. (1992). Is working memory working? The fifteenth Bartlett lecture. *Quarterly Journal of Experimental Psychology*, **44A**, 1-31.

BADDELEY, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology*, **49**, 5-28.

BISANZ, J., & LEFEVRE, J.-A (1990). Strategic and nonstrategic processing in the development of mathematical cognition. In D. F. Bjorklund (Ed.), *Children's strategies: Contemporary views of cognitive development*. Hillsdale, NJ: Erlbaum.

CAMPBELL, J. I. D., & TARLING, D. P. M. (1996). Retrieval processes in arithmetic production and verification. *Memory & Cognition*, **24**, 156-172.

CAMPBELL, J. I. D., & TIMM, J. C. (2000). Adults' strategy choices for simple addition: Effects of retrieval interference. *Psychonomic Bulletin & Review*, **7**, 692-699.

CASE, R. (1985). *Intellectual development: Birth to adulthood*. San Diego: Academic Press.

CRAIK, F. I. M., GOVONI, R., NAVEH-BENJAMIN, M., & ANDERSON, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, **125**, 159-180.

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.

DE RAMMELAERE, S., STUYVEN, E., & VANDIERENDONCK, A. (2001). Verifying simple arithmetic sums and products: Are the phonological loop and the central executive involved? *Memory & Cognition*, **29**, 267-273.

ENGLE, R. W., TUHOLSKI, S. W., LAUGHLIN, J. E., & CONWAY, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology*, **128**, 309-331.

FÜRST, A. J., & HITCH, G. J. (2000). Separate roles for executive and phonological components of working memory in mental arithmetic. *Memory & Cognition*, **28**, 774-782.

GEARY, D. C., & WIDAMAN, K. F. (1992). Numerical cognition: On the convergence of componential and psychometric models. *Intelligence*, **16**, 47-80.

GEARY, D. C., & WILEY, J. G. (1991). Cognitive addition: Strategy choice and speed of processing differences in young and elderly adults. *Psychology & Aging*, **6**, 474-483.

GILHOOLY, K. J., LOGIE, R. H., WETHERICK, N. E., & WYNN, V. (1993). Working memory and strategies in syllogistic-reasoning tasks. *Memory & Cognition*, **21**, 115-124.

GROEN, G. J., & PARKMAN, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review*, **79**, 329-343.

HECHT, S. A. (1999). Individual solution processes while solving addition and multiplication math facts in adults. *Memory & Cognition*, **27**, 1097-1107.

HECHT, S. A., TORGESEN, J. K., WAGNER, R. K., & RASHOTTE, C. A. (2001). The relations between phonological processing abilities and emerging individual differences in mathematical computation skills: A longitudinal study from second- to fifth-grade. *Journal of Experimental Child Psychology*, **79**, 192-227.

HEGARTY, M., SHAH, P., & MIYAKE, A. (2000). Constraints on using the dual-task methodology to specify the degree of central executive involvement in cognitive tasks. *Memory & Cognition*, **28**, 376-385.

KAYE, D. B., DEWINSTANLEY, P., CHEN, Q., & BONNEFIL, V. (1989). Development of efficient arithmetic computation. *Journal of Educational Psychology*, **81**, 467-480.

KIRK, E. P., & ASHCRAFT, M. H. (2001). The perils and promise of using verbal reports to study math strategies. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 157-175.

KRUEGER, L. E., & HALLFORD, E. W. (1984). Why 2 + 2 = 5 looks so wrong: On the odd–even rule in sum verification. *Memory & Cognition*, **12**, 171-180.

LEFEVRE, J.-A, BISANZ, J., DALEY, K. E., BUFFONE, L., & SADESKY, G. S. (1996). Multiple solution routes to solution of single-digit multiplication problems. *Journal of Experimental Psychology: General*, **125**, 284-306.

LEFEVRE, J.-A, BISANZ, J., & MRKONJIC, L. (1988). Cognitive arithmetic: Evidence for obligatory activation of arithmetic facts. *Memory & Cognition*, **16**, 45-53.

LEFEVRE, J.-A, SADESKY, G. S., & BISANZ, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, & 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.

LOGIE, R. H., & BADDELEY, A. D. (1987). Cognitive processes in counting. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 310-326.

LOGIE, R. H., DELLA SALA, S., LAIACONA, M., CHALMERS, P., & WYNN, V. (1996). Group aggregates and individual reliability: The case of verbal short-term memory. *Memory & Cognition*, **24**, 305-321.

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

LORCH, R. F., & MEYERS, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 149-157.

MARSH, R. L., & HICKS, J. L. (1998). Event-based prospective memory and executive control of working memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 336-349.

MCNAMARA, D. S., & SCOTT, J. L. (2001). Working memory capacity and strategy use. *Memory & Cognition*, **29**, 10-17.

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. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, **116**, 250-264.

SIEGLER, R. S., & SHIPLEY, E. (1995). Variation, selection, and cognitive change. In G. Halford & T. Simon (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31-76). Hillsdale, NJ: Erlbaum.

TEASDALE, J. D., DRITSCHEL, B. H., TAYLOR, M. J., PROCTOR, L., LLOYD, C. A., NIMMO-SMITH, I., & BADDELEY, A. D. (1995). Stimulus-independent thought depends on central executive resources. *Memory & Cognition*, **23**, 551-559.

TORGESEN, J. K. (1996). A model of memory from an information processing perspective: The special case of phonological memory. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function* (pp. 157-184). Baltimore: Brookes.

TORGESEN, J. K., & HOUCK, G. (1980). Processing deficiencies in learning disabled children who perform poorly on the digit span task. *Journal of Educational Psychology*, **72**, 141-160.

WOODCOCK, R. W., & JOHNSON, M. B. (1990). *Woodcock–Johnson Test of Achievement Standard Battery, Form B*. Itasca, IL: Riverside.

ZBRODOFF, N. J., & LOGAN, G. D. (1986). On the autonomy of mental processes: A case study of arithmetic. *Journal of Experimental Psychology: General*, **115**, 118-130.