Title

Why Do The Math? The Impact of Calculator Use on Participants' Actual and Perceived Retention of Arithmetic Facts

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 30(30)

ISSN

1069-7977

Authors

Pyke, Aryn
LeFevre, Jo-Anne
Isaacs, Ruby

Publication Date

2008

Peer reviewed

Why Do The Math? The Impact of Calculator Use on Participants' Actual and Perceived Retention of Arithmetic Facts

Aryn Pyke (apyke@connect.carleton.ca)
Institute of Cognitive Science, Carleton University, 1125 Colonel By Drive
Ottawa, ON K1S 5B6 Canada

Jo-Anne LeFevre (jo-anne_lefevre@carleton.ca)
Institute of Cognitive Science and Department of Psychology, Carleton University
Ottawa, ON K1S 5B6 Canada

Ruby Isaacs (risaacs@connect.carleton.ca)
Department of Psychology, Carleton University, 1125 Colonel By Drive
Ottawa, ON K1S 5B6 Canada

Abstract

What is the impact of calculator use on the acquisition of arithmetic facts? Some, but not all, prior research reports that mental practice promotes better subsequent performance than calculator practice (i.e., the *generation effect*). Is answer production faster and more accurate on a test after practice with versus without a calculator? If so, to what extent does mental practice promote retention of the fact, enabling retrieval (semantic memory) versus streamlined computation algorithms (procedural memory)? To investigate this issue, 32 participants practiced sets of 6 problems (3 large, 3 small) 36 times each, either with or without a calculator. Then, in the test phase, participants produced answers to practiced as well as novel problems, without a calculator. Practice without a calculator led to faster, more accurate responses on the test than practice with a calculator. The data further suggest this speed advantage after no-calculator practice was due to retrieval of the facts (e.g., no problem-size effect, many retrieval reports) rather than optimized computation. Interestingly, participants subjectively reported a comparable increase in the proportion of facts memorized over the course of practice with and without a calculator, but fewer retrievals were reported on the actual test after calculator practice, and a substantial problem-size effect remained on response times. Some theoretical and pedagogical implications are discussed.

Keywords: math cognition, calculator, alphabet arithmetic, generation effect, retrieval, education, problem size

Introduction

What are the consequences of the pervasive integration of calculators into the mathematical curricula? One possibility is that calculator access may engender in students more positive attitudes about mathematics and themselves (Roberts, 1980). Another is that the use of calculators for tedious computations may free students' attention and allow them to focus on important conceptual issues. The goal of the present research, however, is to investigate the potential impact of calculator use at a more basic level: the acquisition of simple arithmetic facts (i.e., 5 + 3 = 8).

Regardless of how the answer to a problem is obtained (i.e., with or without a calculator), both components of the arithmetic fact (the problem and answer) become available to the student for association in memory. Thus, repeated practice at solving a problem, with or without a calculator, could promote the long term retention of that arithmetic fact. However, research has provided evidence for a *generation effect* such that subsequent answer production can sometimes be facilitated when students learned by generating the answer themselves, rather than by reading/copying it (McNamara & Healy, 1995) or obtaining it with a calculator (McNamara, 1995).

However, this notion of a generation effect has been applied broadly to various performance enhancements (e.g., Jacoby, 1978, Slameck & Graf, 1978) that may be rooted in quite different memory representations or processes (episodic vs. semantic vs. procedural). In our study, we tested the hypothesis that practice without a calculator would facilitate the retention of the arithmetic fact in *semantic memory* – allowing the participant to subsequently retrieve rather than compute the answer. Alternately, practice without a calculator might optimize computation procedures. We manipulated problem size and included novel problems on the test to assess the relative impacts of practice on semantic and procedural memory.

Results from prior studies on calculator use are ambiguous: answer generation did not always improve performance (for reviews: McNamara & Healey, 1995; Roberts, 1980). Further, researchers often could not control for pre-experimental practice, or precisely manipulate and equate the amount of practice across conditions (calculator vs. no calculator). For our stimuli, we used alphabet-arithmetic facts (e.g., A+4=E; Logan & Klapp, 1991) because our participants had no pre-experimental exposure to these problems, but could readily compute the answers by counting through the alphabet. A laptop was customized to function as a calculator for these problems.

Method

Participants

Undergraduates (N=32, 16 female) received course credit for their participation.

Materials

The stimuli were alphabet arithmetic problems (A + 4 =). To manipulate problem size there were two addends (+2, +4). The alphabet was partitioned into 4 disjoint sets of 3 letters each (i.e., A-C; G-I; M-O; and T-V). Combining each of the 3 letters in a set with each of the 2 addends (+2, +4) produced 6 possible alphabet arithmetic problems per set, for a total of 24 different problems. Note that answers for problems in one set could never also be answers to problems in another set. Each participant saw all four sets of stimulus problems: one set during calculator practice, one set during no-calculator practice, and the other two sets provided the novel problems, respectively, for either the test after calculator practice or the test after no-calculator practice. The role of each stimulus set was counterbalanced across participants and conditions.

Procedure

The experiment was implemented in E-prime and executed on a PC equipped with a microphone and mouse, to record participant responses. The experiment was divided into two conditions, and each included both a practice phase and a testing phase. For each participant, the practice phase in one condition was done with a calculator, and the practice phase in the other condition was performed without a calculator. The order of these two practice conditions was counterbalanced across participants. Each type of practice was followed by a test without a calculator. The entire experimental session lasted for about 90 minutes.

Practice Phase. During practice, the 6 problems in the practice set were each presented 36 times. The presentation order of the six problems was randomized within each of these 36 cycles. In each trial, a prompt (*) appeared in the centre of the screen for 250 ms, then the problem (e.g., G + 4 =) was presented in the centre of the screen, in black font against a white background. Participants were instructed to obtain the answer to the problem as quickly and accurately as possible (either with or without a calculator, depending on the condition). Participants were required to state their answers aloud, which triggered the computer's microphone and caused their response time to be recorded. The experimenter then recorded the participant's response (or an error code in the rare cases in which the microphone did not register the response, or accidentally triggered prior to the response). As soon as a participant had uttered his or her response, the whole fact (G + 4 = K) was then displayed for 1500ms to facilitate forming the problem-answer association in memory. Even if the participant's response was incorrect, the display always presented the correct fact in order to discourage the formation of false associations during training. Participants were told to look at the displayed fact for feedback on their response.

In the calculator practice condition, participants were required to obtain the answer to each problem using a 'calculator' that was implemented on a laptop computer running a customized program in Python. To obtain the answer to a problem, such as "G + 4 =", the participant first pressed the laptop key corresponding to the first operand (G), and this operand appeared on the laptop display. Next, the participant typed the key corresponding to the second operand (4) from among the number keys along the top row of the QWERTY keyboard. Label stickers were placed on these numeric keys to indicate that they expressed the addition operator as well as the addend ("+4"). Then the answer (K) appeared in the centre of the laptop display, in green font against a black background, and the participant stated the answer aloud as quickly as possible. Thus, the calculator was optimized so that a participant only pressed two keys to obtain the answer to a given problem, rather than having to press 4 keys as would be required on a typical arithmetic calculator (first operand, operator, second operand, equal sign). This two-press functionality compensated for the fact that the letter keys are more numerous and spatially distributed than the numeric keys on a conventional calculator, and also helped to minimize motor fatigue, frustration and delay during the 216-trial practice phase (6 problems * 36 cycles). Note that during calculator training, participants were required to use the calculator on each trial, even if they felt they already knew the answer to the problem from memory.

After providing their verbal response on each trial, participants were prompted to indicate whether they felt they knew that answer from memory (trial-end retrieval report). Responses were made using the mouse (left button = not memorized, right button = memorized). Additionally, at the end of the practice phase, participants estimated how many of the 6 practice facts they had committed to memory.

Testing Phase. During testing, no calculator was available so trials were similar to those in the no-calculator practice phase. Participants were tested on the 6 familiar problems from the corresponding practice phase, as well as on 6 novel problems, to provide baseline performance measures. Prior to each test, participants were notified that they might be asked to solve new problems, not previously seen during practice. Problems (practice set and novel set) were each presented twice during the test (24 trials per test).

Results

Our primary interest was in the performance in the test phase: response time, accuracy, and trial-end reports of retrieval. However, for completeness, analyses were also conducted for these three dependent variables on the data from the practice phase. A few trials (<5%) were excluded from the analyses because the participant's response did not trigger the microphone, or another sound prematurely triggered it. Additionally, analyses of response times are based on trials in which correct responses were received.

Practice Phase

For each practice type (with or without a calculator), a full practice session consisted of 36 cycles through a set of 6 problems. For analysis purposes, the 36 practice cycles were partitioned into 6 practice blocks of 6 cycles each. A

2(practice type: calculator vs. no-calculator) x 2(problem size: +2 vs. +4) x 6(practice block: one to six) repeated measures ANOVA was performed for each of the dependent variables: accuracy scores, response times, and trial-end retrieval reports. When the practice block variable failed Mauchly's test of Sphericity, the Greenhouse-Geisser correction was used to assess significance. Any pairwise comparisons were made using the Bonferroni adjustment.

Accuracy Score in Practice Phase. Participants almost always made correct responses when using a calculator to obtain their answers (99.4%), however, they made a few mistakes when answers were obtained without a calculator (92.6%), $F(1,31) = 63.2$, $p = .000$. There was also a main effect of practice block on accuracy, $F(5,82.4) = 5.0$, $p = .005$, and an interaction of practice block with practice type, $F(5,79.4) = 5.7$, $p = .002$. In particular, the calculator always operated at ceiling-level accuracy (aside from random typing errors by the participants), leaving no room for improvement with practice (99.6% in practice block 1 vs. 99.4% in practice block 6, $p = 1.000$), however accuracy did improve with practice in the no-calculator condition (87.6% in practice block 1 vs. 94.8% in practice block 6, $p = .106$). Finally, there was also a main effect of problem size on accuracy, $F(1,31) = 8.0$, $p = .008$, and an interaction between problem size and practice type (calculator vs. no-calculator), $F(1,31) = 5.0$, $p = .032$. In particular, responses were more accurate for smaller problems (+2, 96.7%) than for larger problems (+4, 95.4%), but again this difference was driven by the no-calculator condition ($p = .015$). Unsurprisingly, the calculator was equally adept at the small and large problems ($p = .158$). There was no interaction between practice block and problem size, $p > .05$.

Response Times in Practice Phase. Response time data are shown in Figure 1 over the course of practice for no-calculator practice (top panel) and calculator practice (bottom panel). As suggested by the downward slopes, response times decreased over the course of the practice phase, $F(1,31) = 116.0$, $p = .000$. Thus, when participants practiced with the calculator, they become more adept at operating it with practice, and when they practiced without the calculator they became faster at computation of answers. There was no overall effect of practice type (i.e., calculator, no-calculator) on response time, $F(1,31) = 0.3$, $p = .610$. However, there was an interaction between practice type and practice block, $F(1,31) = 13.4$, $p = .000$. In particular, during the first practice block, participants were faster at obtaining their answers with the calculator than without it ($p = .017$). However, this difference disappears in the middle of the practice session (practice blocks 2 – 4), and by the penultimate and final practice blocks, participants were significantly faster at obtaining answers without a calculator than with it ($p = .022$ for practice block 5, and $p = .017$ for practice block 6). Thus, after 30 exposures these arithmetic problems were faster to compute (or retrieve) mentally than to compute using a calculator.

Participants responded more quickly to small (+2) than to large problems (+4), $F(1,31) = 56.7$, $p = .000$, and there was an interaction of problem size with practice type, $F(1,31) = 51.5$, $p = .000$. In particular, the overall problem size effect is due to the no-calculator condition. The absence of a problem size effect for answers obtained with a calculator ($p = .423$) is unsurprising because participants executed the same number of key presses on the calculator whether the addend was 2 or 4, and the calculator itself is equally fast at producing answers for either addend. In contrast, as seen in Figure 1, the problem size effect in the no-calculator condition diminishes over practice (and is completely eliminated by the test phase), which resulted in a significant interaction between problem size and practice block, $F(1,31) = 7.7$, $p = .000$. The absence of a problem size effect during calculator practice and the diminishing problem size effect during no-calculator practice produced a 3-way interaction of problem size, practice type and practice block, $F(1,31) = 6.7$, $p = .000$.

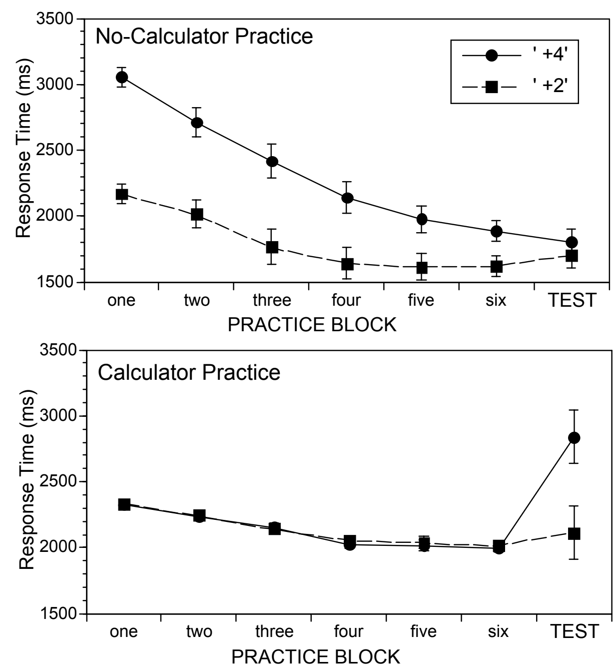

Figure 1. Response times during practice and test phases. Whiskers represent standard error bars for the problem-size comparison in each block.

Trial-End Reports. After each response during the practice phase, participants were asked if they had automatically recalled the answer when the problem was displayed (i.e., did the answer "pop into mind?"). The number of answers reported as automatically retrieved increased over the course of practice, $F(5,94) = 61.3$, $p = .000$ (Greenhouse-Geisser), and reached 71.1% in practice block 6 for the no-calculator condition and 63.5% for the calculator condition. There was no significant effect of practice type, $F(1,31) = 0.5$, $p = .486$: during each of the six blocks of practice, a comparable number of answers were reported as retrieved in the calculator practice condition and the no-calculator practice condition ($ps > .05$). Note, however, that in the calculator practice condition, participants were always required to use the calculator prior to stating their answer, even if they felt they had access to the answer from memory. Thus in the calculator condition, a post-hoc report of retrieval does not reflect how the answer was actually obtained, but rather that the participant felt they had (could have) also retrieved it. Lastly, there was also an effect of problem size, $F(1,31) = 6.2$, $p = .019$, and an interaction of problem size with practice block, $F(5,87.7) = 5.2$, $p = .003$. In particular, during the first half of practice (blocks 1-3), reports of retrieval were more frequent for small versus large problems (*ps* <.05), but by the latter half of practice (blocks 4-6) reports of retrieval were comparable for small and large problems (*ps* >.05).

Practice-End Report. After completing each type of practice (calculator/no-calculator), participants were asked how many of the 6 practice problems they felt they had committed to memory. There was a main effect of practice type: after 36 practice cycles (6 practice blocks of 6 cycles each), participants estimated that they had memorized more problems during practice with a calculator than during practice without a calculator (3.9 vs. 3.3, or 65% vs. 55%, $F(1,31) = 186.2$, $p = .000$). These practice-end estimates are slightly lower than estimates obtained from trial-end reports in the final practice block (71% vs. 64%).

Testing Phase

During testing, no calculator was available. Participants were tested on the set of 6 familiar problems from the preceding practice phase as well as on 6 novel problems. A 2(practice type: calculator vs. no-calculator) X 2(problem size: +2 vs. +4) X 2(problem familiarity: practiced vs. novel) repeated measures ANOVA was performed for each of the three dependent performance measures: accuracy score, response time and trial-end retrieval reports.

***Accuracy Scores in Test Phase*.** There were no main effects of problem size (+2 vs. +4), or problem familiarity (practiced vs. novel), nor were there any significant interactions (*ps* > .10, n=16). However, there was a main effect of practice type: when participants had practiced without a calculator they gave more correct responses on the test than when they had practiced with a calculator (90.5% vs. 86.4%), $F(1,31) = 6.6$, $p = .015$. Note that participants did not have access to a calculator during the test phase itself, so the calculator versus no-calculator contrast in this analysis refers to the type of practice performed during the practice phase immediately before the test in question.

Planned Bonferroni comparisons were conducted to address our key research questions. For example, which type of practice resulted in better accuracy on the test (for those practiced problems)? Participants responded more accurately on the test to problems they had practiced without (vs. with) a calculator ($p = .001$). Which type of practice resulted in better accuracy on the *novel* test problems? Accuracy on novel problems was not affected by practice type ($p = .647$). Thus, practice without a calculator did not result in more accurate computation skill (i.e., for novel problems), but resulted in more accurate production of the answers to practiced problems. Notably, calculator practice provided no increase in accuracy for practiced versus novel (baseline) problems on the test ($p = .707$). In contrast, practice without a calculator provided a 10% improvement in accuracy for practiced relative to novel problems on the test (p =.019).

***Response Times in Test Phase*.** Test responses were faster for small (+2) than for large (+4) problems, $F(1,31) = 83.5$, p =.000, and for practiced problems versus novel problems, $F(1,31) = 48.1$, $p = .000$. Furthermore, test response times were 260 ms faster when the preceding practice phase was conducted without versus with a calculator (2488 ms vs. 2747 ms, respectively), however this main effect of practice type was only marginal in this repeated measures ANOVA, $F(1,31) = 3.0$, $p = .094$. There were, however, significant interactions of practice type with both problem size (+2 vs. +4), $F(1,31) = 5.4$, $p = .027$, and with problem familiarity, $F(1,31) = 15.0$, $p = .001$.

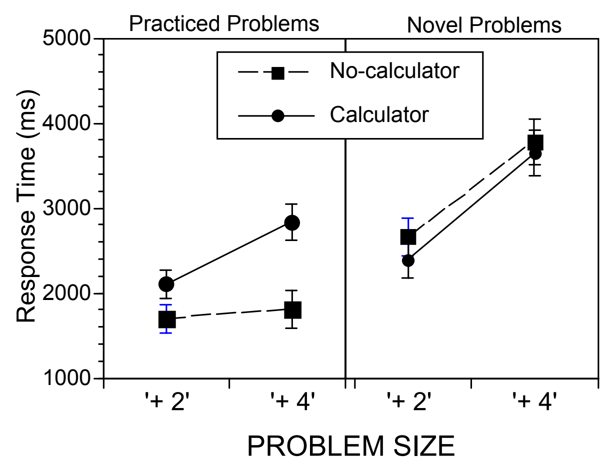

Figure 2. Mean latencies in testing phase for practiced and novel problems by practice condition (no-calculator vs. calculator).

Planned pairwise comparisons using the Bonferroni adjustment provided additional information about these interactions and our key research questions. Figure 2 provides a breakdown of the response times by problem type, problem size and practice type. The left panel of Figure 2 illustrates that test responses were faster for problems practiced without a calculator than for problems practiced with a calculator ($p = .000$). Which type of practice resulted in faster response times on the test for novel problems? As shown in the right panel of Figure 2, for the novel problems in the test, response time was not affected by the type of practice immediately preceding the test ($p = .357$). However, a clear problem size effect was expected for these novel problems because participants had to compute the answers by counting through either 2 or 4 letters. As expected, panel (b) illustrates that there was a significant problem size effect for the novel test problems, regardless of whether the test was preceded by practice with or without a calculator (*ps* = .000). In contrast to novel problems, previously practiced problems may have been committed to memory during practice. Thus, for practiced problems we might not expect a significant problem size effect during testing: regardless of the addend, the participant could simply employ a one-step retrieval process. Was there a problem size effect during testing for the practiced (vs. novel) problems? For problems practiced with a calculator, there remained a significant problem size effect during testing (p = .001), however, for problems practiced without a calculator, there was no significant problem size effect in test response times (p = .304). Thus, in all, the repeated measures analysis outlined above suggests that practice without a calculator did not result in faster computation skill (i.e., for novel problems), but resulted in rapid recall of the answers to practiced problems.

Trial-End Retrieval Reports. After each response during the test, participants reported whether they had retrieved the answer (vs. computed it). Unsurprisingly, participants reported more retrievals for the previously practiced problems than for novel problems on the test (59.0% vs. 16.7%), $F(1,31) = 93.2$, $p = .000$. Furthermore, when participants practiced without a calculator they reported more retrievals during testing than when they had practiced with a calculator (41.1% vs. 34.5%), $F(1,31) = 5.2$, $p = .030$. There was an interaction of practice type and problem familiarity, $F(1,31) = 33.1$, $p = .000$. For practiced problems, when participants had practiced without a calculator they reported more retrievals than when they practiced with one (69.5% vs. 48.4%, $p = .029$), however, for the novel problems, the pattern was reversed: when participants had practiced with a calculator they reported more retrievals than when they had practiced without one (20.6% vs. 12.8%, $p = .000$). Finally, participants reported more retrievals on small versus large problems (41.5% vs. 34.1%), $F(1,31) = 8.0$, $p = .008$. There were no interactions of problem size with problem familiarity or practice type.

Discussion

The present study explored the impact of practice type (calculator vs. no-calculator) on the acquisition and retrieval of arithmetic facts. When problems were practiced without a calculator, participants responded more rapidly and accurately during testing than when problems were practiced with a calculator, despite the fact that each type of practice involved the same number of exposures (36) to each problem. These data are consistent with the generation effect, and, more specifically, with our hypothesis that practice without a calculator facilitates committing facts to memory, so that they can be retrieved rather than computed during subsequent testing.

Another possible explanation for faster test response times after practice without a calculator is that this practice may have enabled participants to become efficient at the computation (counting) process itself. However, there are several pieces of evidence that suggest that, in the present study, the faster response times after practice without a calculator are predominantly due to increased reliance on retrieval rather than optimized computation skill.

First, on a trial-by-trial basis during the *test*, participants reported they had memorized (i.e., retrieved rather than computed answers for) a larger percentage of problems practiced without a calculator than with a calculator. Second, after practice without a calculator, there was no problem size effect in the test response times for those practiced problems. This absence of a problem size effect after practice without a calculator is consistent with the use of a one-step retrieval process[1], regardless of the addend (+2 vs. +4). In contrast, after calculator practice, a significant problem size effect was present in test response times, which is consistent with reliance on a counting process, rather than on one-step retrieval. Finally, practice without a calculator did not facilitate faster computation for the novel problems. That is, practice at the counting process itself did not provide a general advantage, here. That said, our participants may have pre-experimentally attained ceiling-level performance at counting. In general, we expect that practice without a calculator could also contribute to optimized computation.

In all, these data provide converging evidence that practice without a calculator is more conducive to committing facts to memory (resulting in faster and more accurate production during testing) than practice with a calculator. What are the possible practical and pedagogical implications of these findings? Given that calculators are so readily available, is there any practical advantage in committing simple arithmetic facts to memory? Our findings suggest that one advantage is speed: after moderate practice, participants could find the answers faster without versus with a calculator. Thus, answers can be retrieved from memory faster than they can be found using a calculator. In general, the need to find and manually operate a calculator produces a delay in the availability of the answers to arithmetic problems. This delay may result in an increase in cognitive load and confusion when simple facts are required by the individual in the course of solving more complex academic and real-world problems. Thus, despite our ready access to calculators, there may be an advantage to having simple facts committed to memory.

That said, is enforcing practice *without* a calculator strictly necessary for committing a fact to memory, or might enough practice *with a calculator* eventually achieve the same end (i.e., ultimately the individual would no longer resort to the calculator for that problem)? The present research involved 36 practices cycles on a set of 6 distinct problems, and cannot directly speak to the potential effect of additional practice in the calculator condition (i.e., how many practice cycles would be required to eliminate the problem size effect in test response times?). However, the present research does suggest that practice *without* a

[1] Retrieval need not always eradicate the problem-size effect, which may also arise from frequency factors (Ashcraft, 1987; Zbrodoff, 1995), however the +2 and +4 problems were presented with equal frequency in the present research.

calculator is more effective at enabling the participant to commit the fact to memory than practice with a calculator. During testing, participants who had practiced without a calculator reported retrieving (vs. computing) answers for 70% of the familiar problems, whereas those who had practiced with a calculator reported answer retrieval on only 48% of the familiar problems. Thus, while it is possible that extensive practice with a calculator might lead to mental retention of arithmetic facts at asymptote, the present evidence suggests that practice without a calculator is significantly more effective at enabling fact retention and promoting rapid answer retrieval on a production test.

Subjective reports on the percentage of problems retrieved provide important converging evidence for our conclusion that participants memorize more problems after practice without versus with a calculator. However, there were also some unexpected patterns in these subjective report data. For example, there was a problem size effect in the trial-end retrieval reports, such that retrieval was reported more frequently for small problems than for large problems. From an efficiency perspective, there should actually be more incentive to memorize large problems (+4), because they are otherwise more time-consuming and tedious to compute than the small problems (+2). Thus, we might have expected an inverse problem size effect (i.e., more retrievals reported for large vs. small problems).

We suggest two possible explanations for the reverse pattern. First, solving the +2 problems may often have required such little time and effort as to subjectively feel as if the answer was being retrieved rather than computed. Hence, sometimes it may have been difficult for participants to introspectively distinguish between retrieval and computation for these small problems, especially post-hoc, which may have led to false positives in the retrieval reports. Second, participants may have been able to leverage their pre-experimental knowledge of the alphabet when solving small problems: a letter operand might locally prime the two subsequent letters (but not the 4^{th} subsequent letter), thereby facilitating the retrieval of answers for our +2 problems relative to the +4 problems. These suspected influences could likely be eliminated by using larger addends (+4 vs. +8), which might then produce the pattern predicted by efficiency considerations: more frequent retrieval on the large (+8) versus small problems (+4).

Another noteworthy finding in the subjective retrieval reports is that on a trial-by-trial basis during *practice*, calculator users felt that they had memory access to as many of the answers as non-users. However, at test, reported retrieval rates after practice with a calculator were lower than after practice without one. Furthermore, a problem size effect was present in test response times after practice with a calculator, but was absent after practice without a calculator. Thus, it seems that when practicing with a calculator, participants may overestimate their ability to mentally retrieve the answers. During calculator practice, a trial-end report of retrieval does not reflect how the answer was actually obtained (i.e., with a calculator), but rather that the participant felt that they had (also) retrieved the answer. While using a calculator, participants may base their trial-end retrieval reports on the familiarity of the problem and/or of the subsequent answer obtained with the calculator: the post-hoc sense of recognition may have led them to feel that they had (or could have) recalled the answer, inflating their reports of retrieval. In contrast, a report of retrieval during practice without a calculator or on the *test* should be informed by the person's memory of the method they actually had to use to obtain the answer (retrieval vs. counting), and thus provides a more reliable estimate about whether the fact has been internalized.

That said, participants' apparent overestimation of their degree of fact retention during calculator practice raises the possibility that the facts may indeed have been retained, but not in a format conducive to answer *production* (recall). For example, the facts may have been accessible for answer *verification* (recognition). Others have suggested there may be task-specific influences on semantic memory access for production versus verification (e.g., Zbrodoff & Logan, 2000). For our production test, practice without a calculator produced better performance than practice with a calculator. However, we are presently investigating whether practice with and without a calculator could produce comparable performance in a (no-calculator) verification test.

Acknowledgments

Sincere thanks are extended to Jim Davies for the use of his lab equipment for data collection. Support for this research was provided by a grant from NSERC to J. LeFevre.

References

Jacoby, L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *J. of Verbal Learning and Verbal Behavior, 17*, 649–667.

Logan, G. D., & Klapp, S. T. (1991). Automatizing alphabet arithmetic. *JEP:LMC*, *17*, 179–195.

McNamara, D. S. (1995). Effects of prior knowledge on the generation advantage: Calculators versus calculation. *Journal of Educational Psychology, 87,* 307-318.

McNamara D. & Healy A. (1995). A generation advantage for multiplication still training. In A. Healy & L. Bourne, Jr. (Eds.), *Learning and memory of knowledge and skills* (pp. 132-169). Thousand Oaks, CA: Sage.

Roberts, D. M. (1980). The impact of electronic calculators on educational performance. *Review of Educational Research, 50*, 71–98.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *JEP: Human Learning and Memory*, 4, 592–604.

Zbrodoff, N. J. (1995). Why is 9 + 7 harder than 2 + 3? Strength and interference as explanations of the problem-size effect. *Memory & Cognition, 23,* 689-700.

Zbrodoff, N. J. & Logan (2000). When it hurts to be misled: Stroop-type interference in a simple arithmetic production task. *Memory & Cognition, 28*, 1-7.