Journal of Experimental Psychology: General
2001, Vol. 130, No. 2, 299–315

0096-3445/01/$5.00 DOI: 10.1037//0096-3445.130.2.299

Cognitive Arithmetic Across Cultures

Jamie I. D. Campbell and Qilin Xue
University of Saskatchewan

Canadian university students either of Chinese origin (CC) or non-Asian origin (NAC) and Chinese university students educated in Asia (AC) solved simple-arithmetic problems in the 4 basic operations (e.g., 3 + 4, 7 − 3, 3 × 4, 12 ÷ 3) and reported their solution strategies. They also completed a standardized test of more complex multistep arithmetic. For complex arithmetic, ACs outperformed both CCs and NACs. For simple arithmetic, however, ACs and CCs were equal and both performed better than NACs. The superior simple-arithmetic skills of CCs relative to NACs implies that extracurricular culture-specific factors rather than differences in formal education explain the simple-arithmetic advantage for Chinese relative to non-Asian North American adults. NAC's relatively poor simple-arithmetic performance resulted both from less efficient retrieval skills and greater use of procedural strategies. Nonetheless, all 3 groups reported using procedures for the larger simple subtraction and division problems, confirming the importance of procedural knowledge in skilled adults' performance of elementary mathematics.

Knowledge of elementary arithmetic (i.e., simple addition, multiplication, subtraction, and division) is a pervasive requirement of everyday modern life, providing the essential means for dealing with a widely diverse variety of numerical-problem-solving situations. Basic arithmetic also provides the foundation for the more advanced mathematical skills that are central to all modern scientific disciplines. Consequently, understanding this fundamental intellectual skill is an important goal for cognitive science (Ashcraft, 1995; Geary, 1994). In this study, Chinese adults educated in the People's Republic of China and Canadian adults educated in Canada, either of Chinese or non-Asian origin, solved simple arithmetic problems involving the four basic operations (e.g., 3 + 4, 7 − 3, 3 × 4, 12 ÷ 3). Participants also reported their strategy (i.e., direct memory retrieval vs. procedural strategies) after each problem. The purpose was to address three important questions of current research in cognitive arithmetic. First, when adults solve simple-arithmetic problems, what is the relative balance of direct memory retrieval versus use of procedural strategies such as counting or transformation (e.g., 6 + 7 = 6 + 6 + 1 = 13)? Recent evidence suggests that even skilled adults make substantial use of procedures (e.g., LeFevre, Sadesky, & Bisanz, 1996), but no study has attempted to assess this for the entire domain of elementary arithmetic. Second, what determines the problem-size effect (PSE) in cognitive arithmetic? The PSE is the virtually ubiquitous phenomenon that the difficulty of simple-arithmetic problems increases as problem size increases. The PSE has been recognized and studied systematically for over 75 years (e.g., Clapp, 1924), but our study was the first to estimate the relative contributions of retrieval and nonretrieval strategies to the PSE for all four operations. Third, what are the sources of differences in the simple-arithmetic performance of North American and Asian adults? Several studies have found that Asian adults generally outperform North American adults on simple arithmetic (Geary, 1996b; Geary et al., 1997; LeFevre & Liu, 1997). By examining performance as a function of the type of strategy reported, we hoped to identify whether differences in the efficiency of retrieval processes, procedural strategies, or both underlie the overall performance difference.

Direct Retrieval Versus Procedural Strategies in Skilled Performance of Simple Arithmetic

It is widely accepted that young children's performance of arithmetic is often based on counting or other procedural strategies, although some retrieval is evident even during the preschool years, especially for small problems such as 1 + 2 (Siegler & Shrager, 1984). What happens as arithmetic skill develops, however, is more uncertain and controversial (Baroody, 1994). One view is that with accumulating exposure to the basic arithmetic combinations, procedural strategies are gradually replaced by direct memory retrieval. The switch to retrieval is assumed to begin during the early public school grades and to proceed gradually over ensuing years as more and more specific facts are committed to memory (Koshmider & Ashcraft, 1991; Siegler, 1988; Siegler & Shrager, 1984). By the time the typical learner reaches college age, most of the basic arithmetic facts presumably have been encountered so frequently that memory retrieval is the predominant strategy. Indeed, many reaction time (RT) and error phenomena produced by skilled adults performing simple addition and multiplication problems are well explained by retrieval-based models (Ashcraft, 1992; Campbell, 1995; Graham & Campbell, 1992; McCloskey, Harley, & Sokol, 1991).

Although the development of simple arithmetic skill generally proceeds from reliance on procedures to reliance on retrieval, it may be rare to achieve exclusive reliance on direct retrieval. North American university students do not report exclusive reliance on

Jamie I. D. Campbell and Qilin Xue, Department of Psychology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada. We thank Mark Ashcraft for helpful comments on a previous version of this article.

Correspondence concerning this article should be addressed to Jamie I. D. Campbell, Department of Psychology, University of Saskatchewan, 9 Campus Drive, Saskatoon, Saskatchewan S7N 5A5, Canada. Electronic mail may be sent to jamie.campbell@usask.ca.

retrieval for simple addition (Geary & Wiley, 1991; LeFevre, Sadesky, & Bisanz, 1996), multiplication (LeFevre, Bisanz, et al., 1996), division (LeFevre & Morris, 1999), or subtraction (Geary, Frensch, & Wiley, 1993). Geary, Frensch, and Wiley (1993) found, however, that older American adults (61 to 80 years) reported almost exclusive reliance on retrieval (97%). Thus, the substantial use of procedures reported by younger North American adults may be a relatively recent phenomenon.

In contrast to the findings for American samples, university students from China appear to rely almost completely on direct retrieval for the solution of simple addition (Geary, 1996b) and multiplication (LeFevre & Liu, 1997) problems. Little is known, however, about simple division and subtraction of East Asian adults, and we know of no studies to date comparing East Asian and American adult samples on these operations. East Asian samples typically present superior arithmetic performance relative to American samples, but it does not necessarily follow that the former develop exclusive reliance on retrieval for all simple-arithmetic problems. For example, knowledge of the complementary relation between addition and subtraction (e.g., $4 + 3 = 7$ corresponds to $7 - 3 = 4$) and between multiplication and division (e.g., $4 \times 9 = 36$ corresponds to $36 \div 9 = 4$) permits skill in one operation to mediate performance on the complementary operation. Furthermore, such mediation may be very efficient (Campbell, 1997, 1999). In this case, people may not develop direct memory for the majority of simple subtraction and division problems because of the availability of efficient procedural methods. Nonetheless, if we assume that our native Chinese participants represent a group whose educational experiences ought to produce the highest level of performance of simple arithmetic, then this group may be most likely to show exclusive reliance on retrieval if such exclusive reliance commonly develops.

Sources of the Problem-Size Effect

It is probably safe to say that the PSE is the most studied phenomenon in the history of mathematical cognition research (cf. Ashcraft, 1992, 1995; Clapp, 1924; Groen & Parkman, 1972; Norem & Knight, 1930). Virtually every study that has examined effects of problem size for simple arithmetic has found that both RTs and error rates tend to increase with the numerical size of the problem (but see Geary, 1996b, Experiment 1). For adults, correlations of about .6 to .8 are observed between the sum or product of a problem's operands and mean RT for a correct response (Campbell, 1995). Our goal here is to identify sources of the PSE for all four operations.

Three factors potentially contribute to the PSE: (a) Retrieval processes may be more efficient for small-number relative to large-number problems, (b) procedural strategies, although less efficient than retrieval overall, may be more efficient for small-number than for large-number problems, and (c) use of procedures may be less common for small-number than for large-number problems. One reason that use of procedures may be less likely for smaller-number problems is that small-number problems are encountered more frequently and therefore are more likely to have high memory strength relative to large-number problems (Ashcraft & Christy, 1995; Geary, 1996b; Hamann & Ashcraft, 1986). Small-number problems may also be less susceptible to associative *interference*, making memorization and retrieval relatively easy compared with larger-number problems (Campbell, 1995; Siegler, 1988; Zbrodoff, 1995). Thus, when retrieval is used it generally is slower and less accurate for large-number problems. Furthermore, a higher probability of procedure use for large-number problems would produce a PSE because procedural strategies generally are less efficient (i.e., slower and more error prone) than direct retrieval (e.g., LeFevre, Bisanz, et al., 1996; LeFevre, Sadesky, & Bisanz, 1996; Siegler, 1988). Regardless of differential usage of procedures for small-number and large-number problems, however, use of procedural strategies would give rise to a PSE because procedures tend to be more difficult or complex for larger-number problems. For example, a counting strategy for addition generally would involve more incrementing steps for larger-number than for smaller-number problems (LeFevre, Sadesky, & Bisanz, 1996; Siegler & Shrager, 1984).

In summary, lower efficiency of retrieval processes for large-number problems, lower efficiency of procedures for large-number problems, and greater use of procedures for large-number problems may all contribute to the PSE. In principle, the contributions of these three factors to the PSE may vary independently. In the following study, we have obtained strategy reports on a trial-by-trial basis for all the basic problems in each operation. By analyzing the magnitude of the PSE separately for reported retrieval and reported procedure trials, it has been possible to estimate the contribution of retrieval and procedural processes to the PSE in all four basic operations. As we discuss next, these analyses also have permitted us to isolate factors that may contribute to cross-cultural differences in adults' basic calculation skills.

Differences Between Asian and American Adults in Arithmetic Performance

An enormous amount of research has examined cross-cultural differences in early mathematics education and achievement (e.g., Chen & Stevenson, 1995; Stevenson, Chen, & Lee, 1993). One consistent finding is that East Asian children develop numerical and arithmetic skills that are superior to North American children matched for level of education and IQ (e.g., Geary, Bow-Thomas, Fan, & Siegler, 1996; Stevenson et al., 1985). This advantage is not found for other cognitive domains (e.g., spatial cognition) or academic domains (e.g., science; Beaton et al., 1996). The arithmetic advantage for East Asian children probably reflects differences both in the formal education system (e.g., greater concentration on mathematics instruction; Stigler, Lee, & Stevenson, 1987) and in cultural or social factors, such as attitudes toward mathematics achievement and the responsibilities of students and parents to skill development (Chen & Stevenson, 1989, 1995; Geary, Bow-Thomas, et al., 1996).

There has been relatively little research, however, examining adults' cognitive processes in arithmetic across different sociocultural settings. Several studies have found that the arithmetic performance advantage of East Asian children relative to North American children is also observed in young adults' performance (e.g., Geary, 1996b; Geary, Salthouse, Chen, & Fan, 1996; LeFevre & Liu, 1997). Geary (1996b, Experiment 2) compared Chinese and American university students' performance of simple addition. Relative to the American group, the Chinese group presented faster RTs and a smaller PSE and reported exclusive reliance on direct retrieval for simple addition problems. The American group reported direct retrieval for only 73% of addition problems. LeFevre and Liu (1997) compared the simple multiplication performance of

North American (mostly Canadian) and Chinese adults attending college in Canada. The Chinese adults solved multiplication problems more quickly and accurately and produced a much smaller PSE. In contrast to these findings with young adults, Geary et al. (1997) found that the arithmetic competencies of older Chinese and American adults (60 to 80 years old) did not differ. This suggests that the cross-cultural differences in arithmetic performance observed in children and young adults may be a recent development and reinforces the conclusion that the achievement gap is due to cultural (i.e., educational and other social influences) rather than biological or linguistic factors.

In this study, we investigated further the sources of cross-cultural differences in young adults' simple-arithmetic performance. Specifically, we addressed three questions. First, what are the sources of the performance differences? As discussed previously in connection with the PSE, differences among the groups may be attributed to three sources: (a) differential usage of retrieval versus procedural strategies for small-number and large-number problems, (b) size-related differences in the efficiency of retrieval processes, and (c) size-related differences in the efficiency of procedures. The present study permitted us to estimate the contributions of these factors to performance differences among the groups. Second, do these differences extend equally across all four operations, or are the cross-cultural differences operation dependent? Third, whereas Chinese adults educated in China typically outperform North Americans educated in the United States or Canada, are the same differences observed between Chinese adults educated in Canada and Chinese adults educated in China? If cultural factors other than formal education (e.g., parental expectations or positive attitudes toward math) contribute substantially to acquisition of arithmetic skills, then arithmetic performance of Chinese Canadians may be superior to that of non-Asian Canadians but similar to that of native Chinese adults (cf. Chen & Stevenson, 1995).

The Current Study

In the current study, the main analyses of the simple-arithmetic data were based on the so-called standard set of problems (LeFevre, Sadesky, & Bisanz, 1996), which excludes problems involving 0 or 1 as an operand or answer. The 0 and 1 problems are often analyzed separately because many of these items may be solved by general rules (i.e., $x + 0 = x$, $x - 0 = x$, $x - x = 0$, $x \times 0 = 0$, $0 \div x = 0$, $x \times 1 = x$, $x \div x = 1$, $x \div 1 = x$). Here, we tested zero and one problems in all four operations intermixed with problems from the standard set. Previous research had indicated that the rule-governed zero and one problems are solved by retrieving the appropriate rule rather than retrieving a specific fact (Sokol, McCloskey, Cohen, & Aliminosa, 1991). These items therefore provided an opportunity to test whether the groups differed with respect to knowledge and execution of the rules for zero and one problems.

After each arithmetic trial, participants indicated the strategy they had used by selecting from a list that included *transform, count, remember,* and *other* as options (cf. Campbell & Timm, 2000). These options corresponded to categories of simple-arithmetic strategies identified in previous research using a full-report method in which participants described their specific activities after each trial (LeFevre, Bisanz, et al., 1996; LeFevre, Sadesky, & Bisanz, 1996). Note that the rules for solving most 0 and 1 problems are unique to these items. We did not include a *rule* strategy option because explaining these rule-based strategies would amount to instructing participants to select *rule* for zero and one problems. A general concern about strategy reports is whether collection of strategy information changes the way people perform arithmetic (Kirk & Ashcraft, 2001). Using a strategy selection method essentially identical to the current study, Campbell and Timm (2000) compared performance between groups ($ns = 50$) who did or did not provide strategy reports for 144 multiplication or division problems. There was no evidence that collecting the strategy reports affected performance in any substantial way.

We also measured performance on a more complex paper-and-pencil arithmetic test: the French kit arithmetic fluency subtests (Ekstrom, French, & Harman, 1976). Arithmetic subtests of the French kit had been used in several cross-cultural comparisons of arithmetic ability (e.g., Geary et al., 1997; Geary, Salthouse, et al., 1996; LeFevre & Liu, 1997). The subtests we used comprise pure blocks of complex multistep addition or division problems and mixed blocks of multistep subtraction and multiplication problems. The French kit, in addition to measuring simple arithmetic skills, entails a substantial component involving more complex procedural skills and working memory management (e.g., carrying, borrowing, or place keeping; Geary & Widaman, 1992). This allowed us to determine if group differences on the simple arithmetic task generalize to more complicated mental arithmetic problems. Thus, in summary, we measured performance with respect to four types of arithmetic: (a) use and efficiency of retrieval strategies for simple arithmetic, (b) use and efficiency of procedural strategies for simple arithmetic, (c) efficiency of rule-based retrieval for zero or one problems, and (d) relatively complex multistep arithmetic, as measured by the French kit.

Method

Participants

Seventy-two students registered in undergraduate or graduate programs at the University of Saskatchewan participated in this experiment. Participants were recruited from advertisements posted on the university campus and were tested more or less in the order in which they contacted the experimenter. Twenty-four of the students (12 men and 12 women) were non-Asian Canadian (NAC) who had been born in Canada and had completed their elementary and secondary education within Canada. Their mean age was 22 years. Another 24 (12 men and 12 women) were Chinese Canadian (CC). Ten were born either in Hong Kong or in mainland China (average age at immigration was 5 years 1 month), and 14 were born in Canada. All of this group had completed elementary and secondary school in Canada. Their mean age was 22 years. All of this group claimed they were fluently bilingual (English and Cantonese). A third group of 24 students (12 men and 12 women) were Asian Chinese (AC), studying at the University of Saskatchewan on student visas. They all had completed their elementary and secondary education in China. Their mean age was 28 years.

The majority of the AC group were graduate students (19 of 24), whereas the two Canadian groups comprised mainly undergraduates. Given that basic arithmetic skills are acquired and consolidated primarily during elementary education and that all the Chinese university students experienced similar selection pressures before university entrance, it is unlikely that there were significant differences between Chinese undergraduate and graduate students in basic arithmetic competency. For similar reasons, even though our AC group generally had better foreign language skills relative to their peers in China (i.e., due to the English or French language proficiency requirements for admission to Canadian universities), they were likely to be equivalent in basic arithmetic skills (cf. LeFevre & Liu, 1997).

Procedure

Each participant was tested in a quiet lab room in a single session lasting about 1.5 h. The experimenter explained that the study concerned processing of numerical information and that several tasks would be administered in the following order: an arithmetic fluency test (the French kit); a block of 41 number-naming trials; four blocks of about 55 arithmetic trials that included addition, subtraction, multiplication, and division; another block of number naming; four more blocks of simple arithmetic; and then a final block of number-naming trials. LeFevre and Liu (1997) found that Asian Chinese university students studying in Canada reported much less preuniversity calculator use than Canadian university students. To pursue this in connection with our samples, we ended the session with a questionnaire in which we inquired about calculator use: "How often did you use a calculator when doing arithmetic problems (e.g., 65 + 34, 23 × 17)?" Participants provided two ratings on a scale ranging from 1 (*never*) to 5 (*always*), one relating the question to their experiences during elementary school and one relating the question to their experiences during secondary school.

Administration of the French kit required about 15 min. Participants were given a brief break after the second set of naming trials. For naming and simple arithmetic trials, NAC and CC participants responded in English, and AC participants responded in Mandarin. Instructions and examples were given in English to all three groups. In the case of difficulty of a Chinese participant understanding English instructions, Chinese explanations were also provided by the experimenter. No feedback about performance was provided. Participants received $10 for the experiment.

Naming and arithmetic stimuli were presented on two high-resolution monitors using an IBM personal computer with one monitor viewed by the experimenter and the other by the participant. Stimuli were presented horizontally as white characters against a dark background. The participant wore a lapel microphone that activated a relay switch connected to the computer's serial port. The sound-activated relay controlled a software clock accurate to 1 ms.

French Kit

Before the number-naming and simple-arithmetic trials, all participants completed the three arithmetic subtests of the French kit (Ekstrom et al., 1976). The three subtests consist of two pages of multistep addition, followed by two pages of division problems, followed by two pages of mixed sets of multistep subtraction and multiplication. Each page contained six rows of 10 problems. Overall, there were 120 addition, 120 division, 60 subtraction, and 60 multiplication problems. For the subtraction–multiplication test, subtraction and multiplication problems alternated across rows. Each participant was given 2 min per page to solve the problems as quickly and accurately as possible.

In the addition subtest, the participant added three 1- or 2-digit numbers, and for the division task, the participant divided 2- or 3-digit numbers by single-digit numbers. For the mixed subtraction and multiplication subtest, the participant had to subtract 2-digit numbers from 2-digit numbers and multiply 2-digit numbers by single-digit numbers. Instructions for the French kit were presented in English to all participants. We might expect some facilitative transfer of practice from the French kit to the simple-arithmetic task (Charness & Campbell, 1988; but see Frensch & Geary, 1993), but such transfer would be practically equivalent for the three groups.

Number-Naming Task

We tested number naming to estimate contributions of answer production times to arithmetic RTs. The 41 naming stimuli used in the experiment included all single- and double-digit integers ranging from 0 to 20 inclusively plus the products >20 for the simple multiplication facts through 0 × 0 to 9 × 9. All these numbers were presented in Arabic form to each participant in each of three randomized blocks of naming stimuli. The first block was placed immediately before the arithmetic trials, the second after the fourth block of arithmetic trials, and the third at the end of the arithmetic trials.

In each block of naming trials, the 41 numbers to be named were presented individually in a random order on the computer monitor. The participant was asked to name each number as quickly as possible. For each trial, a fixation dot appeared at the center of the screen for 1 s and then flashed twice over a 1-s interval. The number appeared at the fixation point on what would have been the third flash. Timing began when the number appeared and was stopped when the sound-activated relay was triggered by the response. When the response was detected, the number disappeared immediately and was replaced by the fixation dot for the next trial.

Simple-Arithmetic Task

Stimuli and design. The four basic arithmetic operations (addition, subtraction, multiplication, and division) were tested. Addition problems were composed of pairs of numbers between 0 and 9 (i.e., 0 + 0 through 9 + 9). There are 55 possible pairings of the numbers 0 through 9 when commuted pairs (e.g., 5 + 8 and 8 + 5) are counted as one problem. The set of 55 includes 45 nontie problems (e.g., 4 + 7 and 9 + 6) and 10 tie problems (e.g., 0 + 0 and 9 + 9). The corresponding subtraction problems include 45 nontie (e.g., 15 − 7 and 9 − 4) and 10 tie problems (e.g., 2 − 1 and 16 − 8). Similarly, the multiplication problems include 45 nontie (e.g., 5 × 6 and 9 × 4) and 10 tie problems (e.g., 2 × 2 and 8 × 8), and the corresponding division problems include 41 nontie problems (e.g., 12 ÷ 3 and 72 ÷ 8) and 9 tie problems (e.g., 16 ÷ 4 and 49 ÷ 7). The total number of arithmetic stimuli was 429. All problems were presented in Arabic format.

Participants received two sets of four blocks of arithmetic trials. Each set included one block for each of the four operations. Each block included all the tie problems and one order of all the nontie problems in a given operation. Approximately half of the nonties were randomly selected to be tested in one order in the first set of blocks (e.g., 4 + 8, 9 × 4, 16 − 7, 56 ÷ 8), and the reverse order was tested in the second set (e.g., 8 + 4, 4 × 9, 16 − 9, 56 ÷ 7). Thus, each order of each nontie was tested once in the experiment, and each tie was tested twice. The order of problems in each block was independently randomized for each participant. Within each group, 1 participant was assigned to each of the 24 possible orderings for the four operations across blocks. The same sequence of operations was used for the first and second sets of four blocks.

Procedure. Participants were told to respond to each problem both quickly and accurately and, after each answer, to report the strategy used. The experimenter read through the following instructions regarding strategy reports, and the participant kept a copy of the strategy report instructions for reference during arithmetic trials:

> In this study, you are asked to indicate the strategy (procedure) you used to solve each arithmetic problem immediately upon solving it. Studies have indicated that people are consciously or unconsciously using various strategies to solve arithmetic problems. Four different strategies (including unclassified strategies) are proposed in the present experiment: Transform, Count, Remember, and Other. Following are descriptions of each corresponding strategy:
>
> Transform: You solve the problem by referring to related operations or by deriving the answer from some known facts. Examples are "4 + 7 equals 11 because it could be transformed into 7 + 3 + 1," "15 − 7 = 8 because 7 + 8 = 15," and "72 ÷ 9 = 8 because 8 × 9 = 72."
>
> Count: You solve the problem by just counting a certain number of times to get the answer. Please note that here counting is defined as

one by one strict simple counting. Examples are "I counted in my head to get the answer" or "I got the answer by counting silently."

Remember: You solve the problem by just remembering or knowing the answer directly from memory. Here are some examples falling into this category: "I just remembered the answer 15 after I got the problem 6 + 9," "The answer 45 just sort of jumps into my head as I see the problem 5 × 9," or "I just knew the answer 6 as I saw the problem 36 ÷ 6."

Other: You solve the problem by a strategy unlisted here, or you do not know what strategy that you used to solve the problem. Examples are "I do not know," "I cannot figure out the exact strategy I used," or "I guessed the answer."

These four categories will be presented on screen after each problem. Please select the most appropriate category and report it to the experimenter. Please note that if you realized that you made an error, even an operation error (e.g., treat 3 + 3 as 3 × 3), still try to report the procedure you used to get the wrong answer. If you have any questions, please ask the experimenter for help.

Once the instructions had been read, the experimenter explained that after each arithmetic trial, the participant would see a screen with two highlighted words, *Strategy Choices,* appearing at the center and the words *Transform, Count, Remember,* and *Other* below the highlighted words. The four words were separated by six character spaces and always appeared in the same order. The participant was encouraged to first reflect back on the mental processes that occurred during the arithmetic trial and then pick out a corresponding strategy on the screen. The experimenter emphasized that the presented strategies were not meant to encourage use of a particular strategy.

The experimenter pressed the space key to initiate each block of trials, and a message on the computer screen at the beginning of each block indicated which operation would be tested. Before each arithmetic trial, a fixation dot appeared and flashed once over a 1-s interval at the center of the screen. The stimuli appeared on what would have been the third flash with the operation sign (+, −, ×, ÷) at the fixation point. Timing began when the stimulus appeared and ended when the response triggered the sound-activated relay. When a sound was detected, the problem display was cleared immediately, which allowed the experimenter to mark RTs spoiled by failures of the voice-activated relay. The experimenter recorded the stated answer (pressing the *enter* key in the case of a correct response), which typically required <1 s. Once the arithmetic answer was entered, the computer displayed the strategy choice prompts. The experimenter pressed a key to record the strategy choice, which also cleared the screen and displayed the fixation dot for the next trial.

Results

Complex Arithmetic: French Kit Performance

For each participant, we computed a single fluency score equal to the total number of problems answered correctly across the three subtests of the French kit (i.e., addition, division, subtraction–multiplication). We conducted a one-way analysis of variance (ANOVA) on the fluency scores, with group as a between-subjects variable. The mean scores were AC $M = 177$, $SD = 45.6$; CC $M = 133$, $SD = 37.7$; NAC $M = 112$, $SD = 35.1$; $F(2, 69) = 16.64$, $MSE = 1{,}579.14$, $p < .01$. The AC group outperformed the NAC group by 58%, $t(46) = 5.51$, $SE = 11.8$, $p < .001$, and the CC group by 33%, $t(46) = 3.65$, $SE = 12.1$, $p = .001$. The relatively smaller advantage (19%) for the CC group compared with the NAC group approached significance, $t(46) = 1.97$, $SE = 10.5$, $p = .055$. Within the CC group, French kit scores did not differ between those who were born in Canada ($M = 139.5$, $SD = 37.7$, $n = 14$) and those born in China ($M = 123.2$, $SD = 37.5$, $n = 10$), $t(22) = 1.09$, $SE = 15.6$, $p = .307$. With regard to error rates (computed as percentages of total problems completed), the three groups were comparable, with NAC making 7.9% errors, CC 7.7%, and AC 6.5%, $F(2, 69) = 0.63$, $MSE = 21.18$. Thus French kit scores were not contaminated by different speed–accuracy trade-off criteria among the three groups.[1]

On average, AC had several more years of education relative to CC and NAC (i.e., more AC than CC or NAC were graduate students), but it is unlikely that this contributed substantively to their better French kit performance. LeFevre and Liu (1997) compared Asian Chinese and Canadian university graduate students (i.e., matched for education level) on the addition, multiplication, and subtraction subtests of the French kit (i.e., excluding the division subtest). The mean numbers correct were 120.6 and 91.5 for the Chinese and Canadian samples, respectively, $t(36) = 4.29$, $SE = 6.8$. We observed very similar results in our study for the corresponding subtests, with means of 121.7 and 85.6 for AC and NAC, respectively, $t(46) = 4.37$, $SE = 8.25$. Thus, the arithmetic ability differences between our AC and NAC groups were comparable to previous research (see also Geary et al., 1997; Geary, Salthouse, et al., 1996) and do not appear to have been substantively affected by differences in years of education. In the next sections, we examine simple-arithmetic performance in detail, to determine if group differences on the French kit generalize to the most elementary of arithmetic problems. First, however, we present analyses of the naming task.

Number Naming

The purpose of the naming task was to estimate contributions of answer production times to the PSE on RT in the arithmetic task. Consequently, we divided the naming stimuli into sets corresponding to the answers to the small-number and large-number problems in addition and multiplication. A problem was small if $m \times n \leq 25$; otherwise the problem was large (cf. Campbell, 1997; Campbell, Kanz, & Xue, 1999; see analyses of arithmetic task below). For addition, the small-number sums corresponded to the numbers less than or equal to 11, and the large-number sums ranged from 11 to 18. For multiplication, products for the small problems were less than or equal to 25, and products for the large problems ranged from 27 to 81. Table 1 contains mean naming RTs for the small and large answer sets as a function of operation and group.

We conducted an ANOVA on naming RTs, with group as a between-subjects variable and operation and answer size as within-subject variables. AC produced slower RTs (546 ms) on average than CC (506 ms) and NAC (504 ms), $F(2, 69) = 5.29$, $MSE = 10{,}630.40$, $p < .01$. Overall, participants were slower to

[1] Comparisons among the subtests of the French kit are of limited value because the tests are not matched with respect to difficulty and were tested in a constant sequence. For completeness, the mean percentages of French kit items correctly completed for addition were AC = 36.5, CC = 34.3, NAC = 31.3; for division, AC = 45.9, CC = 26.1, NAC = 22.0; for multiplication, AC = 58.4, CC = 40.9, NAC = 31.3; and for subtraction, AC = 71.3, CC = 59.4, NAC = 48.9. Geary et al. (1997), in a comparison of young Chinese and U.S. adults (12th graders), similarly observed a greater Chinese advantage for complex subtraction than for complex addition.

Table 1
Mean Naming RT for the Sets of Small and Large Answers for Addition and Multiplication by the Three Groups

Operation	Answer set S	L	All	L − S
AC				
Addition	540	551	546	11
Multiplication	539	552	546	13
M	540	552	546	12
CC				
Addition	502	501	502	−1
Multiplication	500	519	510	19
M	501	510	506	9
NAC				
Addition	502	501	502	−1
Multiplication	502	510	506	8
M	502	506	504	4

Note. RT is in milliseconds. RT = reaction time; S = small answers; L = large answers; AC = Asian Chinese; CC = Chinese Canadian; NAC = non-Asian Canadian.

name the large answers (523 ms) than the small answers (514 ms), $F(1, 69) = 13.46$, $MSE = 222.64$, $p < .01$. There was some evidence that answers for multiplication were named slightly slower (521 ms) than those for addition (517 ms), $F(1, 69) = 3.28$, $MSE = 114.42$, $p = .07$. No other effects were significant.

AC named addition sums and multiplication products about 40 ms slower than NAC and CC. There was also a slight RT advantage (9 ms) to name the small answers compared with the large answers. Because the small-answer sets involved a substantial number of single-digit numbers, whereas the large-answer sets exclusively involved two-digit numbers, the RT advantage for small answers could be due to encoding processes, production processes, or both. Most important, the size effect in naming was small and did not interact with group or operation. Thus, differences in answer production times did not complicate interpretation of the PSE in the simple-arithmetic task.

Simple Arithmetic

The main analyses of the simple-arithmetic data were based on the standard set of problems (LeFevre, Sadesky, et al., 1996), which excludes problems involving 0 or 1 as an operand or answer. Separate analyses of zero and one problems are presented in a later section. There are 64 problems in the standard set for each operation. To operationalize problem size, the standard set in each operation was divided into two subsets representing small and large problems. For both addition and multiplication, small problems were defined as those with the product of two operands smaller than or equal to 25, and large problems were defined as those with the product of two operands larger than 25 (Campbell, 1997; Campbell et al., 1999). The small and large problems in subtraction and division were defined on the basis of the inverse relationships between addition and subtraction and between multiplication and division. For example, because both 5 + 8 and 8 + 5 were classified as large problems, both 13 − 5 and 13 − 8 were then defined as large problems. Similarly, because both 3 × 4 and 4 × 3 were classified as small problems, both 12 ÷ 3 and 12 ÷ 4 were correspondingly classified as small problems. On the basis of this classification procedure, the small and large problem sets contained the same number of problems (32) in each operation. Per-problem RT and error data broken down by group, operation, and strategy are available from Jamie I. D. Campbell.

Analyses of the standard set are divided into three parts. The first part deals with all arithmetic trials regardless of the strategies reported, the second reports analyses of reported-retrieval trials only, and the third presents analyses of procedure trials.

All Trials

Percentage Reported Retrieval (Remember)

Table 2 contains the mean percentage use of direct retrieval reported (i.e., selection of *remember* from our strategy list) for small and large problems as a function of group and operation. An ANOVA was performed with group as a between-subjects variable and problem size and operation as within-subject variables (see Table 3 for the source table). In keeping with previous research, reported use of direct retrieval was greater for small problems (90%) than large problems (73%). Reported retrieval also varied across operations with multiplication (98%) the highest, followed by addition (88%), subtraction (72%), and division (69%). Size and operation interacted, however. Specifically, the decreases in reported retrieval for large relative to small problems were −29% for subtraction, −21% for division, −14% for addition, and −1% for multiplication.

With respect to effects of group, NAC reported less use of retrieval overall (72%) than CC (87%) and AC (85%). The Group × Operation interaction was significant, however. As Table 2 shows, all three groups reported almost exclusive reliance on retrieval for multiplication (at least 95%), whereas for addition, subtraction, and division collectively, NAC reported less retrieval (63.5% on average) than CC (84.0%) and AC (80.2%). The three-way interaction of Group × Operation × Size approached significance ($p = .054$). As Table 2 shows, the three groups showed similar PSEs on reported retrieval for multiplication, division, and subtraction, but NAC showed a substantially larger PSE for addition (23%) relative to CC (8%) or AC (11%).

Reaction Time

Mean correct RTs (regardless of strategy reported) for small and large problems as a function of operation and group are included in Table 2. Overall, 1,188 RTs (3.9%) were spoiled due to failures of the sound-activated relay. The 481 RTs (1.6%) more than 3 *SD* from each participant's grand mean for each Operation × Problem Size cell were discarded as outliers. The RT means in Table 2 received a Group × Operation × Problem Size ANOVA (see Table 3 for the source table). As expected, RTs were faster for small problems (854 ms) than large problems (1,092 ms). Mean RT also differed across operations, with addition fastest (849 ms), followed by multiplication (930 ms), subtraction (1,026 ms), and division (1,086 ms). The RT PSE (i.e., mean RT for large minus mean RT for small) was largest for subtraction (336 ms), then division (254 ms), multiplication (209 ms), and addition (151 ms).

With respect to effects of group, NAC produced the slowest RT overall (1,107 ms), followed by CC (917 ms) and AC (893 ms).

Table 2
Percentage of Retrieval Reported, Mean RT, and Error Rate for Small and Large Problems as a Function of Group and Operation for All Trials

	% retrieval reported			Mean correct RT			Mean error %		
Operation	S	L	L − S	S	L	L − S	S	L	L − S
AC									
Addition	97	86	11	735	786	52	2.1	3.0	0.9
Multiplication	100	100	0	806	921	116	2.8	5.6	2.8
Subtraction	93	64	29	806	1,066	261	3.8	9.8	6.0
Division	82	59	23	891	1,134	243	5.9	7.6	1.7
M	93	77	16	809	977	168	3.5	6.5	2.9
CC									
Addition	99	91	8	737	855	119	3.0	4.3	1.3
Multiplication	98	97	1	788	949	162	2.7	8.3	5.7
Subtraction	93	65	27	820	1,129	308	5.9	9.4	3.4
Division	86	70	17	940	1,121	181	5.7	11.9	6.3
M	94	81	13	821	1,014	192	4.3	8.5	4.2
NAC									
Addition	88	64	23	848	1,131	283	3.9	5.3	1.4
Multiplication	98	95	3	884	1,235	351	3.2	10.2	6.9
Subtraction	73	42	31	947	1,386	439	3.1	7.0	3.8
Division	68	46	22	1,045	1,385	339	5.3	6.9	1.6
M	82	62	20	931	1,284	353	4.0	7.4	3.4

Note. RT is in milliseconds. RT = reaction time; S = small problems; L = large problems; L − S = problem-size effect; AC = Asian Chinese; CC = Chinese Canadian; NAC = non-Asian Canadian.

The Group × Size interaction was due mainly to a substantially larger PSE for NAC (353 ms) relative to both CC (192 ms) and AC (168 ms). There were no other significant effects in the RT analysis.

Table 3
Analyses of Variance for Mean Percentage of Retrieval, RT, and Error Rate for All Trials

		F		
Source	*df*	Ret%	Mean correct RT	E%
		Between subjects		
Group (G)	2	11.38**	7.32**	0.47
MSE	69	1,215.83	360,946.49	175.73
		Within subjects: Size		
Size (S)	1	178.86**	132.94**	64.75**
G × S	2	2.66	7.94**	0.75
MSE	69	212.90	61,225.22	27.11
		Within subjects: Operation		
Operation (O)	3	96.52**	44.83**	17.08**
G × O	6	4.70**	0.13	3.15*
MSE	207	282.74	35,103.83	21.43
		Within subjects: Size × Operation		
S × O	3	65.95**	14.27**	5.27**
G × S × O	6	2.10*	1.08	2.32*
MSE	207	76.23	15,365.80	20.33

Note. RT is in milliseconds. RT = reaction time; Ret% = percentage of retrieval reported; E% = mean error percentage.
* $p < .05$. ** $p < .01$.

Error Rate

Table 2 includes the mean percentage of errors for small and large problems as a function of group and operation (see Table 3 for the ANOVA source table). Small problems were less error prone than large problems (4.0% vs. 7.5%), and error rates differed across the operations (7.2% for division, 6.6% for subtraction, 5.5% for multiplication, and 3.6% for addition). The PSE on error rates varied across operations, however (3.2% for division, 4.4% for subtraction, 5.1% for multiplication, and 1.2% for addition).

There was no overall effect of group on error rates ($F < 1$; NAC 5.7%, CC 6.4%, AC 5.1%), but the Group × Operation interaction and the triple interaction of Group × Operation × Problem Size were both significant. An examination of Table 2 shows, however, that there was no simple pattern underlying these effects. Instead, the interactions appeared to arise from a complex pattern of relatively small differences. For example, the PSE for error rates was large in multiplication for NAC relative to the other two groups, relatively large in division for CC, but relatively large in subtraction for AC. Nonetheless, the groups did not differ in overall error rate. This makes it possible to interpret the RT data without concern that the group differences in RT reflect different overall speed–accuracy trade-off criteria.

Summary

The results indicate that procedural strategies played an important role in performance of simple arithmetic for all three groups. CC reported procedure use at rates very similar to AC (13% and 15%, respectively), whereas on average NAC reported substantially more procedure use (28%). The Chinese groups reported almost exclusive retrieval for addition and multiplication facts, but they reported substantial use of procedures for subtraction and division. All three groups reported almost exclusive use of retrieval for simple multiplication. Overall, procedural strategies were reported much more often for large (27%) than for small (10%) problems. The three groups were similar overall with respect to the PSE on procedure use, except that NAC presented a larger PSE on procedure use for addition.

The two Chinese groups were very similar in simple-arithmetic performance. To assess this, we conducted follow-up Group × Operation × Size analyses of percentage retrieval, correct RT, and percentage error that included only the two Chinese groups. The only significant group effects (i.e., AC vs. CC) were relatively weak three-way interactions (*ps* ≥ .04) in the RT and error analyses (see Table 2). None of the main effects or two-way interactions in the three analyses approached conventional significance (*ps* ≥ .12). Thus, the two Chinese groups were remarkably similar overall in simple-arithmetic performance. In contrast, NAC produced substantially longer RTs and larger PSEs compared with CC and AC. For simple arithmetic overall, NAC were 24% slower than AC and 21% slower than CC. There were no overall differences in error rates among the groups. To isolate the sources of the performance differences, we next present separate analyses for retrieval and procedure trials.

Trials Reported as Retrieval (Remember)

Reaction Time

Overall, retrieval was reported for 90% of small problems and 73% of large problems (see Table 2). Mean retrieval RTs for small and large problems as a function of operation and group are presented in Table 4 (see Table 5 for the corresponding ANOVA source table). With procedure trials excluded, the PSE was still robust (114 ms; 940 vs. 826 ms for large and small problems, respectively). Thus, less efficient retrieval for large problems contributed substantially to the PSE. Mean retrieval RTs differed across operations, with addition the fastest (797 ms), followed by subtraction (883 ms), multiplication (910 ms), and division (943 ms). The PSE on retrieval RT varied across operations, however: 66 ms for addition, 122 ms for subtraction, 177 ms for multiplication, and 100 ms for division.

The groups differed in overall retrieval RT: NAC produced a substantially slower overall RT (962 ms) than CC (852 ms) and AC (837 ms). NAC also showed an overall PSE (148 ms) that was larger than those of CC (109 ms) and AC (89 ms). This effect was qualified by the significant three-way interaction of Group × Size × Operation. As can be seen in Table 4, the triple interaction was due primarily to the large retrieval PSEs for NAC in multiplication and division relative to those for CC and AC. In fact, with NAC excluded (i.e., performing the RT analysis including AC and CC only), there were no significant main or interaction effects involving the group factor.

To estimate the contribution of the retrieval PSE to the overall PSE for each group and operation, we expressed the PSE for retrieval trials (L − S in Table 4) as a percentage of the PSE for all trials (L − S in Table 2). These percentages appear in Figure 1. For NAC, 42% of the overall PSE was associated with retrieval trials; the proportions were 57% for CC and 53% for AC. As expected, given the high percentages of retrieval

Table 4
Mean RT and Error Rate for Small Problems and Large Problems as a Function of Group and Operation for Retrieval Trials

	Mean correct RT			E%		
Operation	S	L	L − S	S	L	L − S
AC						
Addition	731	768	37	2.2	2.4	0.2
Multiplication	806	921	116	2.7	5.6	2.9
Subtraction	790	915	125	3.3	7.7	4.4
Division	844	921	78	4.9	4.6	−0.3
M	793	881	89	3.3	5.0	1.8
CC						
Addition	735	810	75	3.0	3.7	0.6
Multiplication	780	929	149	2.5	8.0	5.5
Subtraction	795	919	124	5.1	8.3	3.2
Division	879	966	87	4.7	10.3	5.6
M	797	906	109	3.8	7.6	3.8
NAC						
Addition	826	911	86	3.5	4.2	0.7
Multiplication	879	1,146	267	3.2	10.5	7.2
Subtraction	896	1,000	104	2.1	5.8	3.6
Division	955	1,090	136	4.6	4.8	0.2
M	889	1,036	148	3.3	6.3	3.0

Note. RT is in milliseconds. RT = reaction time; E% = mean error percentage; S = small problems; L = large problems; AC = Asian Chinese; CC = Chinese Canadian; NAC = non-Asian Canadian.

Table 5
Analysis of Variance for Mean RT and Error Rate on Retrieval Trials

		F	
Source	*df*	Mean correct RT	E%
Between subjects			
Group (G)	2	5.00**	0.59
MSE	68	167,492.79	191.87
Within subjects: Size			
Size (S)	1	201.89**	38.04**
G × S	2	5.34**	1.57
MSE	68	9,635.57	30.38
Within subjects: Operation			
Operation (O)	3	48.27**	6.80**
G × O	6	1.71	2.72*
MSE	204	11,977.46	28.58
Within subjects: Size × Operation			
S × O	3	13.50**	7.56**
G × S × O	6	2.81*	2.43*
MSE	204	15,365.80	20.85

Note. RT is in milliseconds. One participant from the NAC group was excluded because of no reported retrieval trials for large division problems. RT = correct reaction time; E% = mean error percentage.
$^{*}p < .05$. $^{**}p < .01$.

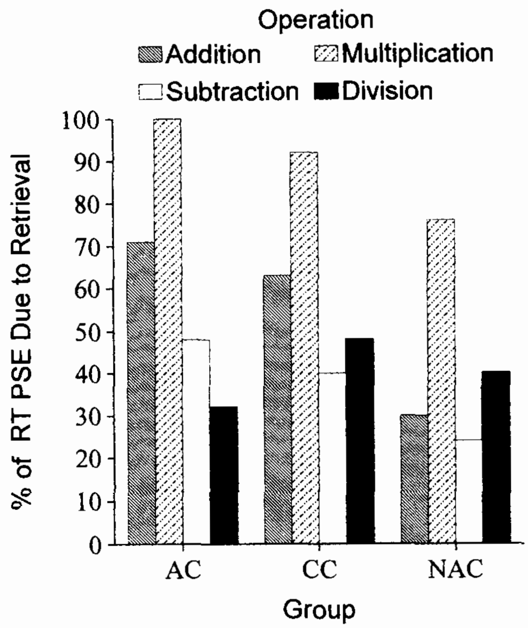

Figure 1. Mean percentage of the reaction time problem-size effect (RT PSE) in the all-trials analysis (large minus small in Table 2) accounted for by the RT PSE for retrieval trials (large minus small in Table 4) as a function of group, operation, and problem size. AC = Asian Chinese; CC = Chinese Canadian; NAC = non-Asian Canadian.

trials reported for multiplication, the retrieval PSE for multiplication accounted for the largest proportion of the overall PSE (76%, 92%, and 100% for NAC, CC, and AC, respectively). For the two Chinese groups, retrieval PSE accounted for the majority of the PSE in addition (over 60%), but for all three groups, retrieval PSE accounted for less than 50% of the overall PSE observed in subtraction and division (from 24% to 48%).

Error Rate

Table 4 includes the error rates for small and large problems as a function of operation and group (see Table 5 for the source table). For retrieval trials, small problems were less error prone than large problems (3.5% versus 6.4%). The error rate for retrieval trials also differed across operations: 3.1% for addition, 5.4% for multiplication, 5.4% for subtraction, and 5.6% for division. The PSE on retrieval errors varied across operations: 0.5% for addition, 5.2% for multiplication, 3.7% for subtraction, and 1.8% for division.

The three groups did not differ in overall error rate for trials reported as retrieval ($F < 1$; NAC = 5.0%, CC = 5.7%, AC = 4.2%). The significant Group × Operation interaction was due mainly to the higher error rates of CC for subtraction and division and to the higher error rate of NAC for multiplication. The three-way Group × Size × Operation interaction was also significant. As Table 4 shows, the interaction was due mainly to the larger PSE on errors for NAC on multiplication and for CC on multiplication and division. Because there were no differences in overall error rates among the three groups, different speed–accuracy trade-off criteria did not contaminate retrieval RT comparisons across groups.

Summary

For all groups and operations, there was a PSE for trials reported as retrieval based, both in terms of RT and errors. Thus, less efficient retrieval processes for large relative to small problems apparently contributed substantially to the PSE. For reported retrieval trials, NAC presented overall longer RTs and a larger PSE compared with the two Chinese groups. The three groups showed substantial variations across operations in the contribution of the retrieval PSE to the overall PSE (see Figure 1). In the following section, we estimate quantitatively the relative contributions of retrieval and procedures to the group differences in RT and the PSE.

Relative Contribution of Retrieval Versus Procedures to Group Reaction Time Differences

A comparison of the RT data in Table 2 and Table 4 provides an estimate of the contributions of retrieval and procedural processes to the group differences in RT and the PSE. We contrasted NAC to the combined results for the two Chinese groups, given the similarity of the RT data for the two Chinese groups. The mean RT for all trials was 1,107 ms for NAC compared with 905 ms for the Chinese groups. In contrast, for retrieval trials, the mean for NAC was 960 ms compared with 845 ms for the Chinese groups. Thus, the RT disadvantages for NAC were 202 ms for all trials and 115 ms for retrieval trials; therefore, procedure trials accounted for 87 ms or 43% of NAC's RT disadvantage.

The PSE for all trials was 353 ms for NAC compared with 180 ms for the Chinese groups (see Table 2). For retrieval trials, the PSE for NAC was 154 ms compared with 99 ms for the Chinese groups. Thus, the PSE for all trials was 173 ms greater for NAC, whereas the PSE for retrieval trials was 55 ms greater for NAC; therefore, procedural trials accounted for 118 ms or 68% of the overall larger PSE for NAC relative to the Chinese groups.

Trials Reported as Procedural Strategies

Rates of Counting, Transformation, and Other

Table 6 contains the percentage of trials reported as transformation, counting, or other procedural strategies as a function of operation and group. Transformation accounted for 90%, 91%, and 95% of procedure trials for NAC, CC, and AC, respectively. We did not record the specific transformation strategies on a trial-by-trial basis, but when participants volunteered details, it was common to report use of addition facts to solve subtraction problems and use of multiplication to solve division problems. NAC reported some use of counting for addition (4.6%) and subtraction (3.8%), whereas the two Chinese groups rarely reported counting for these operations (0.2% and 0.2% for CC, and 0.3% and 0.6% for AC). There was virtually no reported use of counting for multiplication and division problems by any of the three groups. Overall, the *other* category accounted for less than 1% of reported procedural strategies. Because of the overall infrequent use of procedures for addition and multiplication (about 7% of trials overall; see Table 2), we analyzed RTs and error rates on procedure trials for subtraction and division only.

Reaction Time

Table 7 contains mean RTs for small and large division and subtraction problems for the three groups. There were 23 NACs, 16 CCs, and 14 ACs with RTs available for all four cells.[2] Mean RTs for procedures were faster for small problems (1,334 ms) than for large problems (1,564 ms), $F(1, 50) = 47.58$, $MSE = 59{,}332.68$, $p < .01$. We would expect faster procedure RTs for small problems because strategies such as transformation involve smaller (typically faster) mediating number facts for small problems than for large problems. There was no other significant effect (all other Fs < 1). Thus, although the overall RT means varied substantially, the three groups were not statistically different in speed of procedural strategies for simple subtraction and division (AC $M = 1{,}299$ ms, CC $M = 1{,}466$ ms, NAC $M = 1{,}530$ ms).

Error Rate

Procedural strategy error rates for small and large problems for subtraction and division appear in Table 7. There were 23 NACs, 17 CCs, and 15 ACs with procedure trials reported for all four cells. The three groups differed in overall error rate, $F(2, 52) = 3.48$, $MSE = 281.21$, $p < .05$. The mean error rates were AC = 11.7% (SD = 8.4), CC = 14.6% (SD = 11.2), and NAC = 7.7% (SD = 5.4). Post hoc analysis (Tukey's honestly significant difference) confirmed fewer errors for NACs than for CCs, whereas ACs did not differ from NACs or CCs. There was no PSE on errors made on procedure trials ($F < 1$). The Size × Operation interaction approached conventional significance levels, however, $F(1, 52) = 3.18$, $MSE = 170.44$, $p = .08$. As Table 7 shows, the PSE on errors for division (0.4%) tended to be larger than that for subtraction (−4.5%). We presently have no explanation for the trend within procedure trials for small subtraction problems to be more error prone than larger subtraction problems.

Table 6
Percentage Use of Procedural Strategies for the Four Operations by the Three Groups

	Strategy			
Operation	Counting	Transforming	Other	Total
AC				
Addition	0.3	7.3	0.5	8.1
Multiplication	0.0	0.0	0.3	0.3
Subtraction	0.6	19.8	0.8	21.2
Division	0.0	29.3	0.4	29.7
M	0.2	14.1	0.5	14.8
CC				
Addition	0.2	4.4	0.4	5.0
Multiplication	0.0	1.3	1.0	2.3
Subtraction	0.2	19.6	1.2	21.0
Division	0.1	20.3	1.5	21.9
M	0.1	11.4	1.0	12.5
NAC				
Addition	4.6	18.8	0.5	23.9
Multiplication	0.0	3.1	0.6	3.7
Subtraction	3.8	37.9	1.0	42.7
Division	0.0	41.4	1.3	42.7
M	2.1	25.3	0.8	28.2

Note. AC = Asian Chinese; CC = Chinese Canadian; NAC = non-Asian Canadian.

Table 7
Mean RT and Error Rate for Small and Large Subtraction and Division Problems for Procedure Trials

	Mean correct RT			E%		
Operation	S	L	L − S	S	L	L − S
AC						
Subtraction	1,203	1,388	185	19.4	11.0	−8.4
Division	1,215	1,389	174	7.0	9.3	2.3
CC						
Subtraction	1,362	1,543	181	13.8	11.1	−2.7
Division	1,367	1,593	226	13.5	20.2	6.7
NAC						
Subtraction	1,350	1,680	330	7.8	7.1	−0.7
Division	1,384	1,705	321	8.7	7.0	−1.7

Note. RT is in milliseconds. RT = reaction time; E% = mean error percentage; S = small problems; L = large problems; AC = Asian Chinese; CC = Chinese Canadian; NAC = non-Asian Canadian.

Summary

For procedure trials for subtraction and division, the three groups did not differ in overall RT or in the magnitude of the PSE measured in RT. Note that the analyses of procedure RTs were based on relatively small percentages of trials (see Table 6) and fewer participants. Therefore, the analyses had less power to detect group differences relative to the analyses of retrieval trials. Overall, procedure RTs were longer for large than for small problems, but there was no overall PSE on error rates for reported procedure trials. Whereas there were no statistical differences among the groups on procedure RTs, error rates on reported procedure trials did vary among the groups, with NACs presenting the fewest procedural errors. The finding that NACs' RTs were not statistically different from the Chinese groups for procedural strategies while NACs made fewer errors suggests that NACs were equal to or perhaps slightly better than the two Chinese groups with respect to procedure use for simple subtraction and division. Efficiency of procedural strategies (Table 7) apparently did not contribute substantially to the overall subtraction and division advantage of the two Chinese groups (Table 2).

Zero and One Problems

We also examined the efficiency with which the three groups performed zero and one problems in the four operations. All the zero problems can be solved by general rules: addition ($x + 0 =$

[2] More AC and CC participants than NAC participants were excluded due to insufficient procedure-use observations, but there was no evidence that the analyses of procedure RTs involved a subset of substantially lower skilled AC and CC participants. For the 18 AC and CC participants excluded, the mean French kit score was 164.1 (SD = 52.5), and for the 30 included, the mean was 149.2 (SD = 43.4), $t(46) = 1.06$, $SE = 14.0$, $p = .29$. Similarly, with respect to overall arithmetic retrieval RTs (averaged over problem size and operation), the mean RT for excluded participants (811 ms, SD = 136.0) did not differ significantly from those included (864 ms, SD = 130.8), $t(46) = 1.34$, $SE = 39.6$, $p = .19$.

x, $0 + x = x$), subtraction ($x - x = 0$, $x - 0 = x$), multiplication ($x \times 0 = 0$, $0 \times x = 0$), and division ($0 \div x = 0$). For one problems, only multiplication ($x \times 1 = x$, $1 \times x = x$) and division ($x \div x = 1$, $x \div 1 = x$) problems correspond to rules; one addition ($x + 1 = ?$, $1 + x = ?$) and one subtraction ($x - 1 = ?$, $x - m = 1?$) problems must be solved by retrieval of a specific fact or by using a procedure such as counting. We assumed that the *remember* strategy was reported to indicate retrieval of a rule or a specific fact (e.g., $x + 0 = x$, $5 - 4 = 1$). Participants in all three groups reported almost exclusive reliance on retrieval for zero problems (≥96%). The two Chinese groups also reported retrieval almost exclusively for one problems (≥96%), but NACs reported procedures (counting or transforming) 13% of the time for one additions and 17% for one subtractions. Among the zero and one items, only one-addition and one-subtraction problems cannot be solved by a rule, and it was only for these specific problems that participants reported use of counting or transformation strategies.[3] The following analyses examined speed and accuracy of performance separately for zero and one problems. Table 8 contains the mean RTs and error rates as a function of group and operation for zero problems and one problems.

Zero Problems

Reaction time. The three groups differed significantly in overall RT (see Table 9 for source table). NACs had a slower overall RT (883 ms) than CCs (744 ms) and ACs (745 ms). RT differed across operations. The rank order from the fastest to the slowest was addition (707 ms), multiplication (741 ms), subtraction (814 ms), and division (900 ms). The Group × Operation interaction was significant. As Table 8 shows, the interaction was due mainly to the much slower RTs of NACs for subtraction and division.

Table 8
Mean Correct RT and Error Rate for Zero and One Problems as a Function of Group and Operation

	Zero problems		One problems	
Operation	RT	E%	RT	E%
AC				
Addition	722	3.1	737	1.2
Multiplication	726	6.3	695	0.9
Subtraction	736	3.3	753	1.9
Division	796	1.4	784	0.9
M	745	3.5	742	1.2
CC				
Addition	676	1.7	700	2.8
Multiplication	704	3.1	661	0.5
Subtraction	767	5.0	767	1.6
Division	828	0.5	852	3.7
M	744	2.6	745	2.1
NAC				
Addition	724	0.8	778	2.8
Multiplication	793	6.9	735	1.2
Subtraction	940	5.0	910	2.5
Division	1,077	1.4	1,024	5.6
M	883	3.5	862	3.0

Note. RT is in milliseconds. One-addition and one-subtraction problems cannot be solved by general rules. RT = reaction time; E% = mean error percentage; AC = Asian Chinese; CC = Chinese Canadian; NAC = non-Asian Canadian.

Table 9
Analyses of Variance for Mean Correct RT and Error Rate for Zero and One Problems

		F			
		Zero problems		One problems	
Source	*df*	RT	E%	RT	E%
Between subjects					
Group (G)	2	13.13**	0.90	7.21**	3.76*
MSE	69	47,079.11	32.79	62,128.87	20.57
Within subjects: Operation					
Operation (O)	3	40.79**	13.92**	77.90**	6.57*
G × O	6	7.97**	1.96	8.90**	2.31*
MSE	207	12,911.67	21.91	6,430.58	11.93

Note. RT is in milliseconds. RT = reaction time; E% = mean error percentage.
* $p < .05$. ** $p < .01$.

Error rate. The three groups did not differ in overall error rates (see Table 9 for source table). The error rates differed across operations, however. The rank order from the least error prone to the most error prone was division (1.1%), addition (1.9%), subtraction (4.4%), and multiplication (5.4%). The Group × Operation interaction was marginally significant ($p = .07$), but this did not seem to reflect any simple pattern in the data.

One Problems

Reaction time. The three groups differed in overall RT for one problems (see Table 9 for source table). NACs had a slower overall RT (862 ms) than CCs (745 ms) and ACs (742 ms). RT differed across operations. The rank order from the fastest to the slowest was multiplication (697 ms), addition (738 ms), subtraction (810 ms), and division (887 ms). The Group × Operation interaction was significant. As Table 8 shows, the interaction owed mainly to the much slower RTs of NACs for subtraction and division.

Error rate. The three groups differed in overall error rates (see Table 9 for source table). The error rate for NACs was 3.0% compared with 2.1% for CCs and 1.2% for ACs. The error rates differed across operations. The rank order from the least error prone to the most error prone was multiplication (0.9%), subtraction (2.0%), addition (2.3%), and division (3.4%). The Group × Operation interaction also was significant. As can be seen from Table 8, this interaction was due mainly to the much higher error rates for NACs and CCs for division.

[3] Because we did not provide a *rule* strategy option, one might ask why our participants chose *remember* rather than *other* for rule-based zero and one problems. One possibility could be that the experience of remembering a rule is similar to the experience of retrieving a specific fact (e.g., no intervening calculations) but contrasts clearly with the experience of performing counting or transformation strategies. This is not a peculiarity of our strategy-selection method. Kirk and Ashcraft (2001), using a full-report method, found that people often did not distinguish between retrieval of rules and retrieval of answers in their strategy reports for zero and one items.

Summary

All three groups reported almost exclusive reliance on retrieval for zero problems. For one problems, the two Chinese groups reported retrieval almost exclusively, but NACs reported substantial (about 15%) counting or transforming for one-addition and one-subtraction problems. Relative to the two Chinese groups, NACs produced longer RTs for zero and one problems, especially for subtraction and division. NACs were also more error prone in solving one problems. In general, performance of zero and one problems was similar to retrieval performance for small-number problems in the standard set. Indeed, the pattern of Group × Operation RT means observed for zero and one problems was very similar to the pattern observed for small-problem retrieval trials (Table 4): The 12 Group × Operation RT means for small problems in Table 4 correlated .88 and .81 (*ps* $< .001$) with the corresponding zero and one RT means in Table 8.

Calculator Use: Relation to Simple and Complex Cognitive Arithmetic

LeFevre and Liu (1997) found that Asian Chinese university students studying in Canada reported much less calculator use than Canadian university students. Only 10% of their Chinese students reported using a calculator before entering a university, whereas 60% of the Canadian sample claimed to use calculators for simple arithmetic well before they entered a university. We wanted to determine if this pattern held for our samples and whether reported calculator use was predictive of arithmetic performance. The scale used here to rate calculator use ranged from 1 (*never*) to 5 (*always*). The mean ratings for frequency of calculator use were AC $M = 1.5$, CC $M = 2.6$, NAC $M = 2.5$, $F(1, 68) = 153.5$, $MSE = 0.39$, $p < .001$. Participants reported using a calculator more often in secondary (2.9) than in elementary (1.6) school, $F(1, 68) = 19.17$, $MSE = 0.86$, $p < .001$, but the increase was greater for the two Canadian groups than for the AC group, $F(2, 68) = 10.18$, $MSE = 0.39$, $p < .001$. The mean ratings for calculator use during elementary school for NACs, CCs, and ACs, respectively, were 1.6, 1.9, 1.2, whereas the respective means for secondary school were 3.4, 3.3, and 1.9.

Use of calculators was reportedly rare in elementary school for all groups, but especially for AC; 83% of ACs reported never using a calculator in elementary school compared with 44% for NACs and CCs collectively. Among CCs and NACs, 81% reported using a calculator at least sometimes in secondary school (48% reported using a calculator often or always). In contrast, only 25% of ACs reported using a calculator at least sometimes in secondary school, and 46% reported never using a calculator in secondary school. In summary, our results indicated that calculator use was relatively rare for ACs before university. In contrast, CCs and NACs were more likely to begin using a calculator in elementary school, and reliance on calculators was substantial by the time they were in secondary school.

Across the 72 participants, reported calculator use (averaged over elementary and secondary school ratings) was negatively correlated to French kit performance, $r(70) = -.33$, $p = .004$. In contrast, with respect to simple arithmetic, calculator use was not reliably predictive of overall mean RT on the simple-arithmetic tasks, $r(70) = .15$, $p = .216$, overall simple-arithmetic error rate, $r(70) = .20$, $p = .096$, or overall percentage use of retrieval for simple arithmetic, $r(70) = -.19$, $p = .118$. Thus, lower calculator use was specifically associated with greater mental proficiency at the relatively complex procedural operations required for the French kit (i.e., carrying, borrowing, or place keeping). French kit scores were negatively related to simple-arithmetic RT, $r(70) = -.73$, $p < .001$; however, as would be expected, French kit performance increased as time to perform the elementary component problems decreased (Geary & Widaman, 1992). In a multiple regression analysis with French kit performance as the dependent measure, calculator use ($\beta = -.29$, $SE = 4.48$, $p = .005$) and simple-arithmetic RT ($\beta = -.70$, $SE = 0.02$, $p < .001$) each accounted for significant variability, $R^2 = .59$, $F(2, 69) = 49.3$, $MSE = 963.39$, $p < .001$.

Discussion

The discussion is organized around the three main issues developed in the introduction. We begin by addressing the usage of retrieval versus procedural strategies in skilled adults' performance of simple arithmetic. Next we discuss the PSE and its sources. Finally, we discuss the sources of the observed differences in arithmetic performance of North American and Chinese adults.

Memory Retrieval Versus Procedural Strategies in Skilled Adults' Simple Arithmetic

Until recently, it was common to assume in cognitive arithmetic research that skilled adults' simple arithmetic was mediated more or less exclusively by direct retrieval from memory (cf. Ashcraft, 1992; Campbell, 1995). One purpose of this study was to examine comprehensively adults' use of retrieval and procedural strategies for simple arithmetic. Our results were consistent with previous research indicating Chinese adults' almost exclusive reliance on retrieval for addition (93%) and multiplication (99%; Geary, 1996b; LeFevre & Liu, 1997). Both ACs and CCs reported a substantially lower rate of retrieval for subtraction and division problems, however, especially for large problems (<70%). Relative to the two Chinese groups, NACs reported less use of retrieval for addition (76%), subtraction (57%), and division (57%) but were similar to the other groups in reporting almost exclusive use of retrieval for multiplication (96%). The greater use of procedures reported by NACs relative to the two Chinese groups was approximately the same for small and large problems, except for addition (i.e., there was no overall Group × Size interaction). Thus, use of procedures by NACs was not, overall, disproportionately greater for the large problems.

There are both similarities and differences between our strategy results and previous findings. With respect to simple addition, the 76% retrieval reported by NACs was similar to previous research testing young North American adults (e.g., 69% in Campbell & Timm, 2000; 73% in Geary, 1996b; 66% in Hecht, 1999; 71% in LeFevre, Sadesky, & Bisanz, 1996). The 57% retrieval reported by NACs for simple subtraction was somewhat lower than the 71% retrieval for subtraction found by Geary, Frensch, and Wiley (1993) for American university students. For simple division, LeFevre and Morris (1999) found 55% reported retrieval for Canadian undergraduates, which closely matched NACs here (57%), but both of these results were much lower than the 90% retrieval found for division by Campbell and Timm (2000). The

96% retrieval for multiplication reported by our NAC sample was higher than in previous research examining North American university students, although the previous results were quite variable (i.e., 87% in Campbell & Timm, 2000; 78% in Hecht, 1999; 85% in LeFevre, Bisanz, et al., 1996; 59% in LeFevre & Morris, 1999). The variability across studies in reported usage of retrieval suggested that strategy reports were influenced by a variety of factors, including population or sampling differences, speed–accuracy criteria, experimental design, and strategy instructions (cf. Campbell & Timm, 2000; Kirk & Ashcraft, 2001; LeFevre & Morris, 1999). We do not assume, therefore, that our strategy results have absolute validity; but we believe, the pattern of differential use of retrieval across groups, operations, and problem size is generally valid. The three groups received the same strategy instructions, and our instructions did not comment on the merits or normative usage of specific strategies or indicate that strategy use might vary systematically with operation or problem size.

Indeed, the pattern of strategy reports and performance data was coherent and consistent with previous research (e.g., LeFevre, Bisanz et al., 1996; LeFevre, Sadesky, et al., 1996). Whereas retrieval was the predominant strategy reported, procedures were reported more often for large than for small problems. This was expected because small problems are encountered more frequently and may generally be easier to memorize (Geary, 1996b). Moreover, RTs were substantially longer for procedure trials than retrieval trials, and this effect was greater for large problems (cf. Tables 4 and 7). This would arise because procedures generally require more (or more difficult) steps for large than for small problems. People may initially attempt retrieval before switching to a procedure (Campbell & Timm, 2000; Siegler & Shrager, 1984), which also would inflate procedure RTs. Procedural strategies were more error prone compared with retrieval (cf. Tables 4 and 7), an effect that would arise from having to perform multistep procedures under relatively high speed pressure. The strategy reports for zero and one problems also reinforce the validity of the reports: Among the zero and one items, only one-addition and one-subtraction problems cannot be solved by remembering a rule, and it was for these specific zero or one problems that participants were most likely to report use of counting or transformation strategies rather than remembering.

Retrieval use varied across operations, with multiplication (98%) the highest, followed by addition (88%), subtraction (72%), and division (69%). Why would retrieval be more common for addition and multiplication relative to subtraction and division? There probably were both structural and experiential influences. Addition and multiplication are commutative operations (i.e., operand order is irrelevant); indeed, transfer studies indicate that commuted pairs (e.g., 3×6 and 6×3) access a common memory node (Campbell, 1999; Rickard & Bourne, 1996). In contrast, division and subtraction are not commutative, and the transfer studies indicate weak or no transfer between order-reversed pairs (e.g., $18 \div 3 = 6$ and $18 \div 6 = 3$). This implies that such items involve separate memory representations. It follows that within the domain of simple arithmetic, there are many more subtraction and division facts to be memorized than there are addition and multiplication facts to be memorized; consequently, we expected more associative interference in memorizing subtraction and division facts. Greater retrieval interference would decrease the relative efficiency of retrieval and promote use of procedures for subtraction and division (Campbell & Timm, 2000; Siegler & Shipley, 1995). Another consideration is that addition is learned before subtraction, and multiplication is learned before division. Consequently, children initially use knowledge of addition and multiplication to solve corresponding subtraction and division problems (Geary, 1994). The asymmetry whereby addition and multiplication facts are used to solve subtraction and division problems, while the reverse is not true, would tend to reinforce and perseverate a memory-strength advantage for addition and multiplication facts relative to subtraction and division. Indeed, in agreement with the strategy reports, analysis of transfer effects indicates that skilled adults rely on knowledge of multiplication to perform division but not vice versa (Campbell, 1997, 1999; LeFevre & Morris, 1999).

Finally, our data also indicated more use of retrieval for multiplication than addition, especially for the larger problems. Because addition is learned first and multiplication may be solved by repeated addition, less retrieval for addition is perhaps somewhat surprising. The explanation may lie in the relative efficiency of retrieval and procedural strategies for addition and multiplication. According to Siegler's adaptive strategy choice model (Siegler & Shipley, 1995), an arithmetic strategy is selected on the basis of knowledge of its efficiency (i.e., speed and probability of success). In this view, less retrieval for addition than multiplication implies an adaptation to the relative efficiency of retrieval versus procedures for addition and multiplication and ought to be reflected in performance. Indeed, as Table 2 shows, addition performance for larger problems was faster and less error prone relative to the corresponding multiplication problems, despite less reported retrieval for the larger addition problems. Greater use of retrieval for multiplication than addition, despite less efficient retrieval for multiplication (see Table 4), would occur if procedures for multiplication generally were even more inefficient relative to procedural strategies for addition. This is plausible. Counting and transformation strategies often entail one or more incrementing (or decrementing) steps. Even for larger addition problems, these incrementing steps may simply involve counting by ones or twos (e.g., $6 + 8$ is $6 + 6 = 12 + [8 - 6 = 2] + 1 = 13 + 1 = 14$), whereas for multiplication the incrementing steps for large problems necessarily involve larger, more difficult additions or subtractions (e.g., 6×8 is $6 \times 6 = 36 + [8 - 6 = 2 \times 6] = 36 + 6 = 42 + 6 = 48$). There were insufficient procedural multiplication trials for a detailed analysis, but the difficulty of procedural strategies for large multiplication problems was evident in NACs' multiplication data: The 3% difference in use of procedures between small (2%) and large (5%) multiplication problems reported by NACs (see Table 2) accounted for 24% of NACs' overall multiplication PSE (see Figure 1). A small increase in use of procedures for large multiplication problems translated into a substantial increase in the PSE, which suggested that procedural strategies for large multiplication problems were very inefficient. In this case, the difficulty of procedural strategies for large multiplication problems could promote more use of retrieval for multiplication relative to addition, even though retrieval for multiplication is less efficient than retrieval for addition.

Sources of the Problem-Size Effect

The PSE has been studied systematically for decades (e.g., Clapp, 1924; Norem & Knight, 1930), but the present study was the first to comprehensively examine the PSE in skilled adults for all four basic operations. The RT PSE was largest for subtraction (336 ms), then division (254 ms), multiplication (209 ms), and addition (151 ms; see Table 2). The classification of trials as involving retrieval or procedures allowed us to isolate three factors contributing to the PSE: (a) the relative speed of retrieval processes for small and large problems, (b) the relative speed of procedural strategies for small and large problems, and (c) the proportions of retrieval versus procedure usage for small and large problems. In the present experiment, the PSE on RT arose from more use of procedural strategies for large problems and lower efficiency of both procedural and retrieval processes for large problems. Overall, the percentage of procedure use reported was higher for large problems (27%) than for small problems (10%; see Table 2). Given that RTs were substantially slower for procedures compared with retrieval (Tables 4 and 7), the higher rate of procedure use for large problems would contribute to the PSE on RT. Furthermore, execution of procedural strategies was slower for large compared with small problems (Table 7). Thus, use of procedures contributed to the PSE both through more frequent use of procedural strategies for large problems and slower execution of procedures for large problems.

Nonetheless, for all groups and operations, there was a PSE for trials reported as retrieval based, both in terms of RT and errors. For retrieval trials only (i.e., with procedure trials excluded), the RT PSEs across operations were 122 ms for subtraction, 100 ms for division, 177 ms for multiplication, and 66 ms for addition. Memory strength for larger problems would be weaker due to lower normative frequency (and therefore less practice) relative to smaller problems. Other retrieval-related influences, such as greater associative interference for larger problems, might also contribute to the PSE on retrieval trials (Campbell, 1995; Geary, 1996b; Zbrodoff, 1995).

In overview, on the basis of the classification of arithmetic trials into retrieval and procedure categories, the present data indicated that the PSE was due to three main factors: greater use of relatively slow procedural strategies for larger problems and slower execution of both procedures and retrieval processes for larger problems. A comparison of the magnitude of the RT PSE for all trials to the RT PSE for retrieval trials only indicated that on average the percentages of the all-trials' PSE due to retrieval were 42% for NACs, 57% for CCs, and 53% for ACs. Thus, roughly speaking, about half of the PSE on RT was due to lower efficiency of retrieval for large problems, and half was due to greater use of slower procedural strategies for large problems.

Sources of Cross-Cultural Differences in Arithmetic Performance

Simple Arithmetic

The two Chinese groups were practically equivalent in all aspects of their simple-arithmetic performance, and both were substantially better than the NAC group. For the standard set of simple-arithmetic problems, NACs were 24% slower overall than ACs and 21% slower than CCs. The slower RTs and greater PSEs produced by NACs reflected more use of relatively slow procedural strategies for both small and large problems and also slower retrieval processes especially for large problems. The analysis of zero and one problems tested whether the groups differed in accessing the relevant rules, facts, or procedures for these items. In general, the pattern of group differences on zero and one problems paralleled that for the results for small problems in the standard set: AC and CC were very similar, and both were faster than NAC. Thus, NACs' relatively inefficient performance was observed consistently across the whole domain of simple arithmetic.

An important implication of our study is that receiving formal mathematics education in China, as opposed to North America, is not necessarily associated with long-term differences in simple-arithmetic performance. In fact, ACs and CCs were practically equivalent in all aspects of their simple-arithmetic performance, despite the fact that ACs had more years of education on average. Because CCs and NACs both received their formal mathematics education in Canada, the simple-arithmetic advantage of CCs over NACs presumably reflects extracurricular influences. Previous research (Stevenson, Chen, & Lee, 1993; Stevenson et al., 1990) has identified extracurricular cultural differences between North America and East Asia that likely contribute to the differences in early mathematical development and achievement. Indeed, mathematics achievement advantages of Asian American compared with Caucasian American high school students are associated with a variety of motivational factors (e.g., parents and peers holding high academic standards, believing that effort mediates success, pursuing extracurricular instruction or practice, having positive attitudes about achievement; Chen & Stevenson, 1995; see also Whang & Hancock, 1994). The present results suggest that these informal cultural-specific factors, rather than differences in formal education, determine the long-term differences in simple-arithmetic achievement levels of Chinese and North American adults (e.g., Geary et al., 1997; Geary, Salthouse, et al., 1996).

Complex Arithmetic

Whereas the two Chinese groups were equivalent on the simple-arithmetic tasks, ACs outperformed both CCs and NACs on the French kit (33% and 58%, respectively). There also was a trend for CCs to score higher than NACs (19%). Because the French kit tested relatively complex mental arithmetic, the advantage of ACs relative to CCs on the French kit, but not on the simple-arithmetic tasks, suggested that ACs were specifically better at executing the relatively complex procedural operations required by the French kit (e.g., carrying, borrowing, or place keeping). This advantage could be related to a better understanding of the conceptual structures that support arithmetic procedures (e.g., Fuson & Kwon, 1992) or other factors that increase the speed of procedural operations. Geary and Widaman (1992) found that speed of carrying in addition and multiplication was strongly related to performance on the French kit. One possibility could be that better performance on the French kit by ACs is tied to calculator use. The two Canadian groups reported more calculator use beginning in elementary school, and this difference was larger for secondary school. Consequently, ACs might have had more practice performing relatively complex arithmetic without the aid of a calculator. This would promote superior paper-and-pencil skills for complex arithmetic operations such as carrying, borrowing, and place keeping.

In contrast, the slightly better performance of CCs relative to NACs on the French kit probably was due to better performance of the simple-arithmetic components. In fact, covarying out mean simple-arithmetic RT (i.e., averaging over problem size and operation) completely eliminated the advantage of CCs relative to

NACs on the French kit, $F(1, 46) < 1$, $MSE = 632.47$.[4] Nonetheless, for procedure trials in the simple-arithmetic task, CCs were actually more error prone (but no faster) than NACs (see Table 7), suggesting that use of procedures for simple arithmetic (e.g., counting or transforming) might not have been as well developed for CCs as for NACs. Thus, CCs were clearly better with respect to retrieval-based performance of simple arithmetic (Table 4), but procedure-based performance (Table 7) was slightly less efficient. A trade-off between a retrieval advantage for CCs relative to NACs, but a small procedural disadvantage, potentially explains why CCs and NACs presented quite similar performances on the French kit despite the substantial advantage CC presented for simple arithmetic.

Other Potential Sources

Apart from cultural differences, what other factors might contribute to cross-cultural differences in basic mathematics performance? One consideration is that the arithmetic differences reflect general differences in cognitive ability or intelligence. Contrary to this, differences between East Asian and North American samples are still found when IQ is controlled and are observed for arithmetic but not for nonarithmetic measures of cognitive or academic performance (Geary, 1996a; Geary, Salthouse, et al., 1996). Another consideration is linguistic experience. Language has complex effects on arithmetic performance (Cocking & Mestre, 1988; LeFevre, 1998), and there is evidence that arithmetic memory is at least partially language based (Campbell et al., 1999; Dehaene, Spelke, Pinal, Stanescu, & Tsivkin, 1999; but see Noël, Robert, & Brysbaert, 1998). Thus, the fact that our two Chinese groups were fluently bilingual could be a factor in their better simple-arithmetic performance. Geary, Cormier, Goggin, Estrada, and Lunn (1993) found, however, that adults' simple addition and multiplication performance was not affected by degree of bilingualism. Another possible linguistic influence is that Mandarin uses a consistent system for constructing number names (e.g., 12 is *ten two* and 53 is *five ten three*), whereas English is more irregular (e.g., the teens words are quite idiosyncratic). These linguistic differences contribute to differences in Chinese- and English-speaking children's strategies for representing quantity (Fuson & Kwon, 1991, 1992; Miura et al., 1994; but see Towse & Saxton, 1997). Nonetheless, whereas linguistic factors apparently confer an advantage for Chinese children's early mathematical development relative to North American children (Miller, Smith, Zhu, & Zhang, 1995), they apparently do not constrain the ultimate level of skill that can be achieved in arithmetic. Geary, Salthouse, et al. (1996) and Geary et al. (1997) found that arithmetic differences between North American and Chinese samples were observed for younger adults but not for older adults (60 to 80 years old) with IQ and years of education controlled. We would expect an advantage for the older Chinese cohort if the Chinese language per se conferred a permanent, intrinsic advantage for simple arithmetic. The fact that no such advantage is observed argues against the possibility that the arithmetic difference for young adults is due to structural linguistic factors.

Summary and Conclusions

Our results provide several important insights into the nature of skilled adults' simple-arithmetic knowledge and sources of cross-national differences in arithmetic performance. For complex arithmetic (i.e., the French kit), AC substantially outperformed both CC and NAC. In contrast, for simple arithmetic, AC and CC were practically equivalent and both performed substantially better than NAC. The results demonstrate that cross-cultural differences in simple and complex arithmetic are dissociable and probably arise from different influences. ACs' superior performance relative to CCs on complex arithmetic, given their equivalent simple-arithmetic performance, implies better developed cognitive skills for complex arithmetic procedures (e.g., borrowing, carrying, place holding). In keeping with this, calculator use was negatively correlated with French kit performance, and ACs reported much less calculator use before college relative to both Canadian groups. ACs evidently had more practice performing relatively complex arithmetic without the aid of a calculator. Interestingly, calculator use was not predictive of simple-arithmetic performance, indicating that greater calculator use is not necessarily associated with weaker elementary arithmetic skills.

Another important conclusion from our experiment is that cross-cultural differences in young adults' simple-arithmetic skills are not directly attributable to cross-national differences in formal education. ACs and CCs, despite receiving their formal math education in China and Canada, respectively, were practically equivalent in all aspects of simple arithmetic. In contrast, CCs and NACs, who both received formal mathematics education in Canada, presented substantively different skill levels. The superior simple-arithmetic performance of CCs relative to NACs implies that extracurricular cultural-specific factors (cf. Chen & Stevenson, 1995; Stevenson, Chen, & Lee, 1993; Stevenson, Lee, et al., 1990) rather than cross-national differences in formal education underlie differences in simple-arithmetic performance observed between Chinese and North American adults. NACs' relatively poor simple-arithmetic performance resulted both from greater use of slow procedural strategies (e.g., counting or decomposition) and from slower retrieval processes especially for numerically larger problems. In contrast, NACs presented equivalent or perhaps slightly better efficiency for procedures for simple arithmetic (i.e., lower error rates relative to CCs). NACs might simply have been more practiced at using such strategies. Nonetheless, NACs' overall greater use of procedures, which are generally less efficient than retrieval, contributed to their overall relatively poor simple-arithmetic performance.

Finally, our study was the first to comprehensively address the role of retrieval versus procedural strategies in skilled adults' basic arithmetic, examining performance of all four basic operations by the same participants. The groups reported very little procedure use for simple multiplication, but NACs reported substantial procedure use for basic addition, and all three groups reported relying substantially on procedural strategies for the larger simple subtraction and division problems. Overall, about half of the PSE in simple arithmetic (i.e., longer RTs for large-number problems relative to small-number problems) was due to slower retrieval for large problems, and half was due to greater use of slower procedural strategies for large problems. The results indicate that procedural strategies constitute a major part of educated adults' repertoire of simple-arithmetic skills. Indeed, exclusive reliance on retrieval for simple arithmetic probably is a rare achievement.

---

[4] Our thanks to an anonymous reviewer for suggesting this analysis.

Even ACs, whose educational and cultural experiences apparently provide an especially strong foundation for simple arithmetic, did not report having memorized all the basic number facts. The findings confirm the central importance of procedural knowledge in skilled adults' performance of elementary mathematics.

References

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44,* 75–106.

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

Ashcraft, M. H., & Christy, K. S. (1995). The frequency of arithmetic facts in elementary texts: Addition and multiplication in Grades 1–6. *Journal for Research in Mathematics Education, 5,* 396–421.

Baroody, A. J. (1994). An evaluation of evidence supporting fact-retrieval models. *Learning and Individual Differences, 6,* 1–36.

Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1996). *Mathematics achievement in the middle school years: IEA's Third International Mathematics and Science Study (TIMSS).* Chestnut Hill, MA: Boston College.

Campbell, J. I. D. (1995). Mechanisms of simple addition and multiplication: A modified network-interference theory and simulation. *Mathematical Cognition, 1,* 121–164.

Campbell, J. I. D. (1997). On the relation between skilled performance of simple division and multiplication. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1140–1159.

Campbell, J. I. D. (1999). Division by multiplication. *Memory & Cognition, 27,* 791–802.

Campbell, J. I. D., Kanz, C. L., & Xue, Q. (1999). Number processing in Chinese–English bilinguals. *Mathematical Cognition, 5,* 1–39.

Campbell, J. I. D., & Timm, J. (2000). Adults' strategy choices for simple addition: Effects of retrieval interference. *Psychonomic Bulletin and Review, 7,* 692–699.

Charness, N., & Campbell, J. I. D. (1988). Acquiring skill at mental calculation in adulthood: A task decomposition. *Journal of Experimental Psychology: General, 117,* 115–129.

Chen, C., & Stevenson, H. W. (1989). Homework: A cross-cultural examination. *Child Development, 60,* 551–561.

Chen, C., & Stevenson, H. W. (1995). Motivation and mathematics achievement: A comparative study of Asian-American, Caucasian-American, and East-Asian high school students. *Child Development, 66,* 1215–1234.

Clapp, F. L. (1924). *The number combinations: Their relative difficulty and frequency of their appearance in textbooks* (Research Bulletin No. 1). Madison, WI: Bureau of Educational Research.

Cocking, R. R., & Mestre, J. P. (Eds.). (1988). *Linguistic and cultural influences on learning mathematics.* Hillsdale, NJ: Erlbaum.

Dehaene, S., Spelke, E., Pinal, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science, 284,* 970–974.

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests: 1976.* Princeton, NJ: Educational Testing Service.

Frensch, P. A., & Geary, D. C. (1993). Effects of practice on component processes in complex mental addition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 433–456.

Fuson, K. C., & Kwon, Y. (1991). Chinese-based regular and European irregular systems of number words: The disadvantages for English-speaking children. In K. Durkin & B. Shire (Eds.), *Language and mathematical education* (pp. 211–216). Milton Keynes, England: Open University Press.

Fuson, K. C., & Kwon, Y. (1992). Korean children's single-digit addition and subtraction: Numbers structured by ten. *Journal for Research in Mathematics Education, 23,* 148–165.

Geary, D. C. (1994). *Children's mathematical development.* Washington, DC: American Psychological Association.

Geary, D. C. (1996a). International differences in mathematical achievement: Their nature, causes, and consequences. *Current Directions in Psychological Science, 5,* 133–137.

Geary, D. C. (1996b). The problem-size effect in mental addition: Developmental and cross-national trends. *Mathematical Cognition, 2,* 63–93.

Geary, D. C., Bow-Thomas, C. C., Fan, L., & Siegler, R. S. (1996). Development of arithmetical competencies in Chinese and American children: Influences of age, language, and schooling. *Child Development, 67,* 2022–2044.

Geary, D. C., Cormier, P., Goggin, J. P., Estrada, P., & Lunn, M. C. E. (1993). Mental arithmetic: A componential analysis of speed-of-processing across monolingual, weak bilingual, and strong bilingual adults. *International Journal of Psychology, 28,* 185–201.

Geary, D. C., Frensch, P. A., & Wiley, J. D. (1993). Simple and complex mental subtraction: Strategy choice and speed-of-processing differences in younger and older adults. *Psychology and Aging, 8,* 242–256.

Geary, D. C., Hamson, C. O., Chen, G.-P., Fan, L., Hoard, M. K., & Salthouse, T. A. (1997). Computational and reasoning in arithmetic: Cross-generational change in China and the United States. *Psychonomic Bulletin and Review, 4,* 425–430.

Geary, D. C., Salthouse, T. A., Chen, G.-P., & Fan, L. (1996). Are East Asian versus American differences in arithmetical ability a recent phenomenon? *Developmental Psychology, 32,* 254–263.

Geary, D. C., & Widaman, K. F. (1992). Numerical cognition: 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 and Aging, 6,* 474–483.

Graham, D. J., & Campbell, J. I. D. (1992). Network interference and number-fact retrieval: Evidence from children's alphaplication. *Canadian Journal of Psychology, 46,* 65–91.

Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review, 79,* 329–343.

Hamann, M. S., & Ashcraft, M. H. (1986). Textbook presentations of the basic addition facts. *Cognition and Instruction, 3,* 173–192.

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

Kirk, E. P., & Ashcraft, M. H. (in press). Do adults overreport strategy use in single digit arithmetic? An investigation of instructional bias in verbal reports. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Koshmider, J. W., & Ashcraft, M. H. (1991). The development of children's mental multiplication skills. *Journal of Experimental Child Psychology, 51,* 53–89.

LeFevre, J. (1998). Interactions among encoding, calculation, and production processes in the multiplication performance of Chinese–English bilinguals. *Mathematical Cognition, 4,* 47–66.

LeFevre, J., Bisanz, J., Daley, K. E., Buffone, L., Greenham, S. L., & Sadesky, G. S. (1996). Multiple routes to solution of single-digit multiplication problems. *Journal of Experimental Psychology: General, 125,* 284–306.

LeFevre, J., & Liu, J. (1997). The role of experience in numerical skill: Multiplication performance in adults from China and Canada. *Mathematical Cognition, 3,* 31–62.

LeFevre, J., & Morris, J. (1999). More on the relation between division and multiplication in simple arithmetic: Evidence for mediation of division solutions via multiplication. *Memory & Cognition, 27,* 803–812.

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

McCloskey, M., Harley, W., & Sokol, S. M. (1991). Models of arithmetic fact retrieval: An evaluation in light of findings from normal and brain-damaged subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 377–397.

Miller, K. F., Smith, C. M., Zhu, J., & Zhang, H. (1995). Preschool origins of cross-national differences in mathematical competence: The role of number-naming systems. *Psychological Science, 6,* 56–60.

Miura, I. T., Okamoto, Y., Kim, C. C., Chang, C.-M., Steere, M., & Fayol, M. (1994). Comparisons of children's cognitive representation of number: China, France, Japan, Korea, Sweden, and the United States. *International Journal of Behavioral Development, 17,* 401–411.

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

Norem, G. M., & Knight, F. B. (1930). The learning of the one hundred multiplication combinations. *Yearbook of the National Society for the Study of Education, Part II,* 551–569.

Rickard, T. C., & Bourne, L. E., Jr. (1996). Some tests of an identical elements model of basic arithmetic skills. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1281–1295.

Siegler, R. S. (1988). Strategy choice procedures and the development of multiplication skills. *Journal of Experimental Psychology: General, 106,* 250–264.

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

Siegler, R. S., & Shrager, J. (1984). A model of strategy choice. In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229–293). Hillsdale, NJ: Erlbaum.

Sokol, S. M., McCloskey, M., Cohen, N. J., & Aliminosa, D. (1991). Cognitive representations and processes in arithmetic: Inferences from the performance of brain-damaged patients. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 355–376.

Stevenson, H. W., Chen, C., & Lee, S. Y. (1993). Mathematics achievement of Chinese, Japanese, and American children: Ten years later. *Science, 259,* 53–58.

Stevenson, H. W., Lee, S.-Y., Chen, C., Stigler, J. W., Hsu, C. C., & Kitamura, S. (1990). Contexts of achievement: A study of American, Chinese, and Japanese children. *Monographs of the Society for Research in Child Development, 55*(1–2, Serial No. 221).

Stevenson, H. W., Stigler, J. W., Lee, S.-Y., Lucker, G. W., Kitamura, S., & Hsu, C. C. (1985). Cognitive performance and academic achievement of Japanese, Chinese, and American children. *Child Development, 56,* 718–734.

Stigler, J. W., Lee, S., & Stevenson, H. W. (1987). Mathematics classrooms in Japan, Taiwan, and the United States. *Child Development, 58,* 1272–1285.

Towse, J. N., & Saxton, M. (1997). Linguistic influences on children's number concepts: Methodological and theoretical considerations. *Journal of Experimental Child Psychology, 66,* 362–375.

Whang, P. A., & Hancock, G. R. (1994). Motivation and mathematics achievement: Comparisons between Asian-American and non-Asian students. *Contemporary Educational Psychology, 19,* 302–322.

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.

Received April 12, 1999
Revision received March 13, 2000
Accepted June 16, 2000 ■