Journal of Cognitive Psychology

ISSN: 2044-5911 (Print) 2044-592X (Online) Journal homepage: https://www.tandfonline.com/loi/pecp21

Retrieval-induced forgetting of arithmetic facts across cultures

Jamie I. D. Campbell, Yalin Chen & Anna J. Maslany

To cite this article: Jamie I. D. Campbell, Yalin Chen & Anna J. Maslany (2013) Retrieval-induced forgetting of arithmetic facts across cultures, Journal of Cognitive Psychology, 25:6, 759-773, DOI: 10.1080/20445911.2013.820191

To link to this article: https://doi.org/10.1080/20445911.2013.820191

Published online: 06 Aug 2013.

Journal of Cognitive Psychology, 2013
Vol. 25, No. 6, 759–773, http://dx.doi.org/10.1080/20445911.2013.820191

Retrieval-induced forgetting of arithmetic facts across cultures

Jamie I. D. Campbell[1], Yalin Chen[2], and Anna J. Maslany[1]

[1]Department of Psychology, University of Saskatchewan, Saskatoon, Canada
[2]Lab of Cognitive Neuroscience and School of Psychology, Nanjing Normal University, Jiangsu, China

Retrieving a single-digit multiplication fact ($3 \times 4 = 12$) can slow response time (RT) for the corresponding addition fact ($3 + 4 = 7$). The present experiment investigated effects of problem type (i.e., tie addition problems such as $3 + 3$ vs. non-ties such as $3 + 4$) and cultural background on this retrieval-induced forgetting (RIF) phenomenon in young adults. Canadians answering in English ($n = 36$), Chinese adults answering in English ($n = 36$), and Chinese answering in Chinese ($n = 36$) received four blocks of multiplication practice and then two blocks of the addition counterparts and control additions. Tie addition problems presented a robust RIF effect that did not differ between groups, but only the Canadian group showed RIF for non-ties and only for small non-ties with sum ≤ 10 ($3 + 4$). The Chinese groups' RIF effect for addition ties, but not small non-ties, converges with recent evidence that ties are solved by direct memory retrieval whereas small non-ties may be solved by highly efficient procedural processes in skilled performers.

Keywords: Arithmetic; Culture; Retrieval-induced forgetting

Retrieval-induced forgetting (RIF) occurs when repeated practice of a memory item impairs retrieval of related, unpractised items. This phenomenon has been demonstrated in numerous memory tests with a wide variety of stimuli (Anderson, 2003; Anderson, Bjork, & Bjork, 1994; Storm & Levy, 2012). Several recent studies have investigated RIF in retrieval of elementary addition (e.g., Campbell & Dowd, 2012; Campbell & Therriault, 2013; Campbell & Thompson, 2012b) and multiplication facts (Galfano et al., 2011; Phenix & Campbell, 2004). In the addition RIF research, participants practise a subset of simple multiplication problems ($2 \times 3 = ?$, $4 \times 4 = ?$, etc.) and then are tested on the addition counterparts (i.e., $2 + 3 = ?$, $4 + 4 = ?$). RIF is expressed in slower response times (RT) or increased errors relative to addition problems whose multiplication counterparts have not been practised. Campbell and Thompson (2012b; see also Campbell & Phenix, 2009; Maslany & Campbell, 2013) showed that the effect occurs following retrieval practice of multiplication facts (repeatedly answering $4 \times 5 = ?$) but not following study practice (reading equations silently and stating the answer, e.g., $4 \times 5 = 20$). The retrieval dependency of RIF in addition favours an inhibitory model of the effect over a model based only on cue-related retrieval interference (Anderson, 2003; Storm & Levy, 2012).

Campbell and Dowd (2012) outlined a model of addition RIF in this paradigm. They proposed that presentation of a target multiplication problem

Correspondence should be addressed to Jamie Campbell, Department of Psychology, University of Saskatchewan, 9 Campus Drive, Saskatoon, SK, Canada S7N 5A5. E-mail: jamie.campbell@usask.ca

This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada. We thank Emily Yan for assistance with data collection.

(e.g., 4×5) primes a family of related facts including the addition counterpart ($4 + 5 = 9$). If a related fact is a strong competitor of the target memory then its priming attracts inhibition during or after target retrieval that can produce a net reduction in competitor accessibility (i.e., RIF). Such competition dependence appears to be a boundary condition for RIF (Anderson, 2003; Levy, McVeigh, Marful, & Anderson, 2007; Norman, Newman, & Detre, 2007; Storm & Levy, 2012; but see also Jakab & Raaijmakers, 2009). In contrast, a related fact that is a weak competitor may be primed, but not inhibited, giving a net increase in accessibility. Whether the transfer effect of multiplication practice on the addition counterpart is negative (RIF) or positive (priming) depends on the balance of inhibition and priming that results from multiplication retrieval practice (Campbell & Dowd, 2012).

The competition dependence of RIF probably explains why the effect is not uniform for all single-digit plus single-digit addition problems. Campbell and Thompson (2012b) tested non-Asian Canadian adults and observed RIF only for small addition problems with sums ≤ 10, but not the larger simple addition problems. Small problems generally have greater memory strength than larger problems (Campbell & Xue, 2001; LeFevre, Sadesky, & Bisanz, 1996; Zbrodoff & Logan, 2005). Given competition dependence, we would expect stronger RIF for small addition problems than large problems because the former are strong competitors to their multiplication counterparts, whereas the latter are not. Campbell and Therriault (2013) provided another illustration of the competition dependency of RIF in arithmetic. They had participants practise rule-governed multiplication problems (e.g., 1×6, 0×4) and multiplication facts (e.g., 2×5, 3×4) for four blocks and then tested the addition counterparts (e.g., $1 + 6$, $0 + 4$, $2 + 5$, $3 + 4$) and control additions. The rule-governed zero and one problems (i.e., $0 \times N = 0$, $1 \times N = N$, $0 + N = N$) are solved by applying a rule rather than by retrieving a specific fact (Campbell & Metcalfe, 2007; Jost, Beinhoff, Hennighausen, & Rösler, 2004; Pesenti, Depoorter, & Seron, 2000). Campbell and Therriault (2013) reasoned that if RIF of arithmetic facts is competition dependent, RIF would not occur for the addition counterparts of the $0 \times N = 0$ and $1 \times N = N$ problems. Rule-based performance of multiplication would not involve retrieval of problem-specific representations, but instead retrieval of a rule-specific procedure that would be applied to all problems that satisfy the conditions of the rule. Consequently, answering $0 \times N$ and $1 \times N$ problems would not generate RIF of specific addition counterparts. Indeed, addition RT and errors increased relative to controls only for problems corresponding to multiplication facts, with no problem-specific effects on addition counterparts resulting from practice of rule-governed multiplications.

Campbell and Therriault (2013) also proposed that the presence of rule-related problems contributed to the RIF results reported by Campbell and Dowd's (2012) examination of arithmetic RIF in adult Chinese-English bilinguals. Participants practised multiplication problems (e.g., $6 \times 7 = ?$), answering a subset in L1 (Chinese) and another subset in L2 (English). Then separate groups answered corresponding addition problems ($6 + 7 = ?$) and control addition problems in either L1 or L2. The large simple addition problems (sum > 10) presented a 75 ms RT cost after their multiplication counterparts were practised in the same language relative to practice in the other language. Unlike Campbell and Thompson (2012b), small addition problems presented no clear evidence of transfer in the Campbell and Dowd (2012) study; however, one third of the small problems used in the Campbell and Dowd study involved 1 as an operand. Inclusion of rule-based items in Campbell and Dowd's small-problem set would have diluted evidence of RIF for small problems. Campbell and Therriault (2013) reanalysed the Campbell and Dowd (2012) data for small problems but excluded the rule-related problems. The results demonstrated a -49 ms RIF effect for the remaining small addition problems when multiplication was practised and addition tested in the same language, but not when practice and test language differed (-9 ms). Thus, both the small and large problem sets tested in the Campbell and Dowd experiment provided evidence of language-specific RIF in Chinese-English bilinguals' memory for addition facts. In contrast, facilitative priming of large addition facts following practice of the corresponding multiplication problems was evident when language switched at test.

THE PRESENT STUDY

The following study examined another potential boundary condition for RIF in arithmetic. Specifically, RIF was examined separately for tie (e.g., $4+4$, $5+5$) and non-tie problems ($3+4$, $4+5$). Recent evidence suggests that ties and non-ties may be processed differently, potentially making them differentially susceptible to RIF. Fayol and Thevenot (2012) proposed that addition ties are mediated by declarative fact retrieval whereas non-tie additions are often mediated by procedures. In two experiments their participants were tested on simple addition and multiplication problems up to 9×9 and $9+9$ using three visual presentation conditions: the two digits and operation sign ($+$ or $\times$) appeared together simultaneously, or the sign preceded or followed the digits by 250 ms. The critical finding was that a preview of the operation sign had little effect on multiplication but facilitated addition performance for all problems with the exception of ties. Fayol and Thevenot (2012) proposed that whereas simple multiplication relies primarily on direct memory retrieval, in contrast, simple addition often relies on highly automatic procedures that can be pre-activated by a preview of the $+$ sign. Such procedures may exist for both small and large non-tie problems, but not ties. Barrouillet and Thevenot (2013) provided converging evidence that different processes mediate addition ties and non-ties. They examined additions involving the numbers 1 through 4, and found that RT increased linearly with problem size (the sum of the operands), and the slope for non-ties (20 ms) was about twice that for the ties (11 ms). They proposed that the small non-ties activated a compacted counting procedure that gives rise to a relatively steep linear problem size effect, whereas the ties are solved by direct fact retrieval from memory (Barrouillet & Thevenot, 2013). Ties may constitute a distinct subcategory of arithmetic problems that are especially well memorised and rely entirely on fact retrieval rather than on procedures (see also Campbell & Gunter, 2002).

The distinction between fact-based and procedure-based arithmetic is directly relevant to RIF. As we explained earlier, Campbell and Therriault (2013) found that rule-governed arithmetic facts involving 0 and 1 did not present RIF, whereas small fact-based additions (that included a mix of tie and non-tie problems) did present RIF. They concluded that performance based on direct retrieval was subject to RIF but procedure-based performance was not. We can relate this to Fayol and Thevenot's (2012) and Barrouillet and Thevenot's (2013) conclusions about non-tie and tie additions; specifically, that the former are often solved by procedures whereas the latter are solved by fact retrieval. In this case, we would expect RIF to be robust for tie problems and weaker or absent for non-tie problems.

All the previous research on RIF in arithmetic has included a mix of ties and non-ties, with no attempt to measure RIF separately for them. Here, that was one of the primary goals of the analyses conducted. The stimuli included all the simple multiplication and corresponding addition problems composed of pairs of digits from 2 through 9, but excluded rule-related multiplication problems (i.e., $0 \times N = 0$ and $1 \times N = N$) and their addition counterparts. Three groups of adult participants were tested: Canadians answering in English (CA), Chinese answering in English (CNe), and Chinese answering in Chinese (CNc). This allowed us to investigate the different RIF effects observed by Campbell and Thompson (2012b) for Canadians compared to Campbell and Dowd (2012) for Chinese participants (e.g., the latter but not the former presented RIF for large addition problems). The stimuli and procedures were somewhat different in these studies, whereas here the groups were tested using identical methods. Thus, these groups provided the opportunity to try to replicate the previous RIF results testing the corresponding groups under comparable conditions (i.e., Campbell & Dowd, 2012; Campbell & Thompson, 2012b), while also exploring potential RIF differences for tie and non-tie problems.

METHOD

Participants[1]

A total of 72 Chinese-English bilinguals were recruited via online advertising at the University of Saskatchewan. They include 39 women and 33 men, aged 17 to 40 years ($M = 24.1$), of which 69 are right handed. Remuneration was \$10 CA or bonus course marks. All reported having been born in China and completing primary and

[1] The experiment was conducted in two phases. A first group of 54 participants (18 per group) were tested in the autumn and winter of 2011/2012. Subsequently, to increase power and precision, a second group of 54 participants was tested in winter 2013. In the interval, the issue of different processes for ties and non-ties emerged (Barrouillet & Thevenot, 2013; Fayol & Thevenot, 2012) and we extended the theoretical framework and analytic approach to apply the experiment to this issue.

secondary education there. Participants were alternately assigned to the L1 (Chinese responding in Chinese; CNc) or L2 (Chinese responding in English; CNe) group. Of those responding in Chinese, 70 used Mandarin and 2 used Cantonese. Mean self-rated proficiency in English on a five-point scale from poor (1) to excellent (5) was 3.25. A third group of 36 non-Asian Canadians (CA) participated for bonus marks in their introductory psychology course at the University of Saskatchewan. The recruitment description of the experiment identified it as a study of simple numerical skills. The sample included 26 women and 10 men, of which 34 are right handed, and all indicated normal or corrected-to-normal vision. Ages ranged from 18 to 26 years ($M = 20.3$). None had participated in previous arithmetic RIF studies.

Apparatus and stimuli

Stimuli appeared on two high-resolution monitors using Eprime 1.0 (Psychology Software Tools, Pittsburgh, PA) with one monitor viewed by the participant and the other by the experimenter. The participant sat approximately 50 cm from their monitor and wore a lapel microphone connected to the computer's serial port through Eprime's SRbox. Sound input through the microphone provided the stop-signal to a software clock accurate to within ± 1 ms.

Addition problems ranged from $2 + 2$ to $9 + 9$ and multiplication from 2×2 through 9×9. All stimuli appeared as black, Courier New, 14-point characters against a white background. Problems were composed of Arabic digits separated by the operation sign ($+$ or $\times$) with adjacent spaces (e.g., 4×9 or $4 + 9$) and displayed in horizontal orientation.

There are 36 possible pairs composed from the operands 2 through 9 when commuted pairs (e.g., $3 + 8$ and $8 + 3$) are counted as one problem. The set includes 8 ties (e.g., $2 + 2$, $3 + 3$, etc.) and 28 non-ties (e.g., $2 + 3$). Non-tie problems appeared with the small number on the left side. The set of 36 operand pairs was divided into two matched sets of 18. The sets were similar to those used by Campbell and Thompson (2012b), but modified to match the sets for the numbers of ties ($n = 4$), small non-tie problems with sums ≤ 10 ($n = 6$) and large non-tie problems ($n = 8$). Set 1 consisted of $2 + 2$, $4 + 4$, $7 + 7$, $9 + 9$, $2 + 5$, $3 + 4$, $2 + 6$, $3 + 5$, $3 + 6$, $2 + 8$, $4 + 7$, $3 + 9$, $4 + 8$, $5 + 7$, $5 + 8$, $6 + 7$, $6 + 8$, $7 + 9$. Set 2 consisted of $3 + 3$, $5 + 5$, $6 + 6$, $8 + 8$, $2 + 3$, $2 + 4$, $2 + 7$, $4 + 5$, $3 + 7$, $4 + 6$, $2 + 9$, $3 + 8$, $5 + 6$, $4 + 9$, $5 + 9$, $6 + 9$, $7 + 8$, $8 + 9$. Multiplication counterparts were created by substituting the times sign ($\times$) for the plus sign ($+$). The addition problem-set corresponding to the unpractised multiplication problems served as control addition problems.

Design

The experiment had two phases. In the practice phase, there were four blocks of the 18 multiplication problems involving either Set 1 or Set 2, with set counterbalanced across participants within each group. The multiplication set practised was the MP set and the set unpractised was the MU set. During the practice phase, participants were instructed to "state the correct answer aloud" for each problem (e.g., for 2×5 say "ten"). The post-practice test phase consisted of two blocks of all 36 addition problems (i.e., both the MP and MU sets). We included a second block to determine if the RIF effect is eliminated (i.e., switched off) by the first retrieval attempt in Block 1 or if it persists and is still measurable in Block 2.

Procedure

The study took place in a quiet room with an experimenter present and required about 45 minutes. Instructions for both multiplication and addition phases requested participants to respond quickly but accurately. Prior to the addition blocks, instructions for strategy reports appeared on the monitor and were read aloud by the experimenter as follows:

> After each problem please indicate how you solved the problem by choosing from among the following strategies: Say TRANSFORM if you used a procedure or knowledge of another related problem. Say COUNT if you used a strategy based on counting one by one. Say REMEMBER if you recalled the answer directly without any intermediate steps. Choose OTHER if you used some other strategy or are uncertain.

Participants also received a sheet of strategy descriptions that included examples of problems not used in the experiment. Numerous studies support the utility of self-reports as experimental measures of adults' strategies for simple arithmetic (e.g., Campbell & Alberts, 2009; Grabner & De Smedt, 2011; LeFevre et al., 1996; Metcalfe & Campbell, 2011; Seyler, Kirk, & Ashcraft, 2003; but see Kirk & Ashcraft, 2001; Thevenot, Castel, Fanget, & Fayol, 2010).

Each trial began with the presentation of a one-second central fixation dot. The problem then appeared with the times sign or plus sign at fixation. Timing began with the onset of the stimulus and continued until the participant's verbal response stopped the timer. When the response (or any sound) was detected, the stimulus was instantly removed from the screen. This allowed the experimenter to identify trials on which the microphone failed to detect response onset and mark them as spoiled. On addition trials in the post-test, participants were prompted to report their strategy immediately after their arithmetic answer. The experimenter recorded the reported strategy with a button press. On all trials, the experimenter also entered the participant's arithmetic response, and then the fixation dot immediately appeared to signal the start of the next trial. Participants received no feedback regarding speed or accuracy.

After completion of the addition-test blocks, all participants completed the three arithmetic subtests of the French Kit (Ekstrom, French, & Harman, 1976). This provided a standardised measure of basic arithmetic skill or fluency. The three subtests consist of two pages of multi-step addition, and end with two pages of division problems, followed by two pages of mixed sets of multi-step subtraction and multiplication. Each page contains 6 rows of 10 problems. Overall, there are 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 two minutes 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.

RESULTS

French-Kit performance

One CA and one CNc participant did not complete the French Kit. The total number of correct answers across the three subtests of the French Kit (addition, division, subtraction/multiplication) was tallied for each of the remaining 106 participants. A one-way ANOVA with group as a between-participants factor confirmed group differences. The mean scores were: CA, $M = 93$, $SE = 5.6$; CNe, $M = 169$, $SE = 5.3$; CNc, $M = 157$, $SE = 5.9$ ($F(2, 103) = 54.4$, $MSE = 1095$, $p < .001$). The CNc and CNe groups outperformed the CA by 69% ($t(68) = 8.0$, $SE = 8.1$, $p < .001$) and 82% ($t(69) = 10.0$, $SE = 7.7$, $p < .001$) respectively. French-Kit performance of the two Chinese groups was similar to that observed by Campbell and Xue (2001) for their Asian-Chinese group (mean score of 177), but the mean score for CA ($M = 93$) was significantly lower than Campbell and Xue's non-Asian Canadian group ($M = 112$, $t(47.4) = 2.14$, $SE = 9.1$, $p = .04$). Thus, the well-established gap in arithmetic skills between Chinese and non-Asian North American university students (e.g., Geary, Salthouse, Chen, & Fan, 1996; Geary et al., 1997; Imbo & LeFevre, 2010) was robust in the present study.

Arithmetic transfer tasks

A total of 537 multiplication practice or addition test RTs (5.2%) were marked by the experimenter as spoiled and thus excluded from analysis. Non-tie problems with a sum ≤ 10 were defined as small and those with a sum > 10 were defined as large (see also, e.g., Campbell & Dowd, 2012; Campbell & Thompson, 2012b; LeFevre et al., 1996). We used the same size definition for addition and multiplication, which provides problem-size results for each operation that are based on the same addition and multiplication counterparts. We analysed median RT rather than mean RT because of the small number of trials per participant in each Group $\times$ Block $\times$ Set (i.e., MP or MU problems) cell; specifically, four ties, six small non-ties, and eight large non-ties. This required that we conduct separate analyses of the three problem types because median RT should not be used to compare conditions with different numbers of trials (Miller, 1988). We first

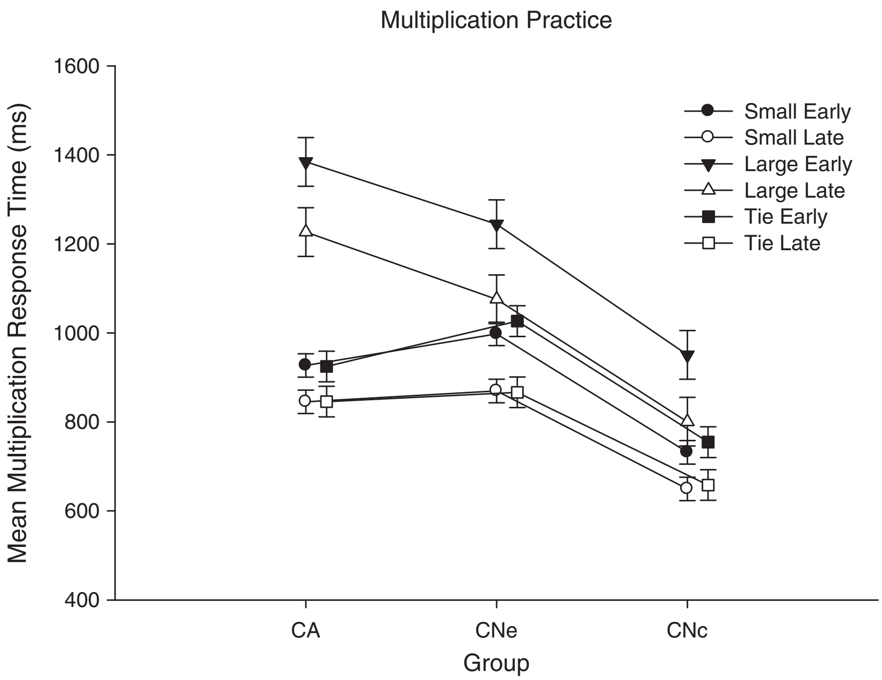

Figure 1. Mean response time in the multiplication practice phase. "Early" refers to practice Blocks 1 and 2, "Late" refers to practice Blocks 3 and 4. Error bars are the 95% repeated-measures confidence intervals based on separate ANOVAs for ties, small, and large problems (Jarmasz & Hollands, 2009; Masson & Loftus, 2003).

present multiplication practice results and then the addition test data to examine potential transfer effects.

Multiplication practice

Multiplication response times

Figure 1 presents mean median multiplication RT for correct answers during multiplication practice by group (CA, CNe, CNc), problem type (ties, small or large), and practice block with pairs of successive blocks combined. We conducted separate Group × Block ANOVAs for each problem type and used the *MSE* for the two-way interaction in each case to compute the 95% repeated-measures confidence intervals (Jarmasz & Hollands, 2009; Masson & Loftus, 2003) for each problem type in Figure 1.[2] The *MSE*s for the Group × Block effect for ties, small and large problems were, respectively, 10,896, 6363, and 27,374. As the important differences in multiplication RT across conditions are apparent given the narrow confidence intervals, we do not report separate inferential statistics for Figure 1 (e.g., see Masson & Loftus, 2003).

Early in practice (i.e., Blocks 1 and 2), large problems produced longer average RT (1193 ms) than small multiplications (886 ms) or ties (902 ms). All three problem types were performed faster later in practice. Specifically, speed up in Blocks 3 and 4 for large, small, and tie problems was, respectively, 159 ms, 97 ms and 111 ms. CNe sped up 152 ms on average from early to late practice, compared to 106 ms for CA and 110 ms for CNc. Figure 1 shows that all groups' RT performance for each of the problem types benefited from practice.

Multiplication error rate

The low average error rate for small non-tie (1.4%) and tie (2.3%) multiplication problems precluded meaningful factorial analysis. For the large non-tie multiplications, a Group × Block ANOVA indicated more errors for CA (10.5%) than CNe (6.0%) and CNc (3.9%) ($F(2, 105) = 6.6$, $p = .002$, $MSE = 124.6$, $\eta_p^2 = .11$). There were more errors early in multiplication practice (7.9%) than late in practice (5.7%) ($F(1, 105) = 7.4$, $p = .008$, $MSE = 33.4$, $\eta_p^2 = .07$). The reduction in errors was 4.7% for CA and 2.3% for CNe, whereas for

[2] Confidence intervals and effect size values were calculated using MorePower 6.0 (Campbell & Thompson, 2012a). MorePower 6.0 is freely available at https://wiki.usask.ca/pages/viewpageattachments.action?pageId=420413544

CNc there was a nominal increase of -0.5% ($F(2, 105) = 3.7$, $p = .03$, $MSE = 33.4$, $\eta_p^2 = .07$).

Participants presented accuracy and/or speed gains during multiplication practice, confirming transfer of practice for the practised problems. The following analyses examined potential item-specific transfer of multiplication practice to performance of the addition counterparts.

Addition test

The dependent variables included addition RT for correct answers, percentage reported memory retrieval, and error rate. Addition transfer effects from multiplication practice are represented by differences between addition MU and MP conditions (i.e., MU − MP). For both RT and errors, negative transfer (RIF) has occurred if MU < MP (i.e., performance is slower or more error prone for MP problems), whereas positive transfer (priming) has occurred if MU > MP (i.e., performance is faster or less error prone for MP problems). Therefore, for RT and errors a negative difference represents RIF and a positive difference represents priming. For percentage of retrieval reported this reverses because when MU < MP this indicates more use of retrieval for MP problems, which implies retrieval priming. When MU > MP in percentage retrieval, retrieval has been inhibited (i.e., RIF) and a procedural strategy promoted. Therefore, for percentage of retrieval a positive difference represents RIF and a negative difference represents priming.

Addition response times

Figure 2 presents mean median RT for each group for tie problems (e.g., $4 + 4$), small non-tie problems (sum ≤ 10) and large non-tie problems (sum > 10). Each problem type received a Group × Block × Set ANOVA. The *MSE* for the three-way interaction in each case was used to compute the 95% CIs for each problem type shown in Figure 2. The *MSE*s for ties, small and large problems were, respectively, 13,642, 5031, and 13,695. Unlike the multiplication RT data in Figure 1, the interpretation of the patterns of means and confidence intervals in Figure 2 is less clear cut. Consequently, we reported the details of the ANOVAs for the addition RT analysis.

Tie addition. Mean RT for ties was faster in Block 2 (806 ms) than in Block 1 (890 ms) ($F(1, 105) = 43.2$, $p < .001$, $MSE = 17{,}563$, $\eta_p^2 = .29$). CNc produced the fastest mean RT for ties (753 ms) followed by CA (864 ms) and CNe (928 ms) ($F(2, 105) = 10.3$, $p < .001$, $MSE = 110{,}081$, $\eta_p^2 = .16$). There were no other significant effects of group in the analysis of tie RT (all $p > .75$).

There was a −39 ms RIF effect for tie additions overall ($F(1, 105) = 8.8$, $p = .004$, $MSE = 3994$, $\eta_p^2 = .08$), but as Figure 2 shows, RIF was present in Block 1 (−70 ms, $t(107) = 3.6$, $p = .0004$, $SE = 19.2$, $\eta_p^2 = .11$) but not in Block 2 (−8 ms) (Block × Set interaction: $F(1, 105) = 7.6$, $p = .007$, $MSE = 13{,}642$, $\eta_p^2 = .07$). Generating an answer to an MP tie addition problem in Block 1 largely eliminated the RIF effect in RT in Block 2. For the two Chinese groups there is some overlap of the 95% confidence intervals for tie MP and MU means in Block 1 (see Figure 2), but a separate analysis of these two groups confirmed a −66 ms RIF effect ($t(71) = -2.9$, $p = .004$, $SE = 22.4$, $\eta_p^2 = .11$).

Small non-tie addition. CNc answered small problems fastest (768 ms) followed by CNe (894 ms) and CA (984 ms) ($F(2, 105) = 12.0$, $p < .001$, $MSE = 141{,}388$, $\eta_p^2 = .19$). Block 1 produced slower RTs (932 ms) than Block 2 (832 ms) ($F(1, 105) = 88.6$, $p < .001$, $MSE = 12{,}189$, $\eta_p^2 = .46$. The Group × Set interaction ($F(2, 105) = 3.3$, $p = .04$, $MSE = 7992$, $\eta_p^2 = .06$) reflected a −38 ms RIF effect for CA ($t(35) = -2.6$, $p = .01$, $SE = 14.9$, $\eta_p^2 = .16$) but not for either CNc (+6.8 ms) or CNe (+2.7 ms). Thus, whereas the two Chinese groups displayed RIF for ties in Block 1 they displayed no RIF for small non-ties. A separate Type (small non-ties vs. ties) × Block × Set analysis for the two Chinese groups confirmed the three-way interaction ($F(1, 71) = 4.9$, $p = .03$, $MSE = 7937$, $\eta_p^2 = .07$). In the corresponding analysis for CA, the only effect of set was the main effect, indicating an overall RIF effect of −44 ms ($F(1, 35) = 8.3$, $p = .007$, $MSE = 17{,}071$, $\eta_p^2 = .19$) with no evidence of a Block × Set interaction ($F(1,35) = 1.1$, $p = .30$, $MSE = 14{,}124$, $\eta_p^2 = .03$). Therefore, in contrast to the Chinese groups, there was no evidence for CA that RIF differed between ties and small non-ties, or that it was diminished in Block 2 compared to Block 1.[3]

[3] Comparisons between problem types should be interpreted cautiously because the median RTs for ties and small non-ties were based on different numbers of trials, which can bias comparisons of median RT (Miller, 1988). If the bias is constant across other experimental factors, however, interaction effects with problem type would be bias free.

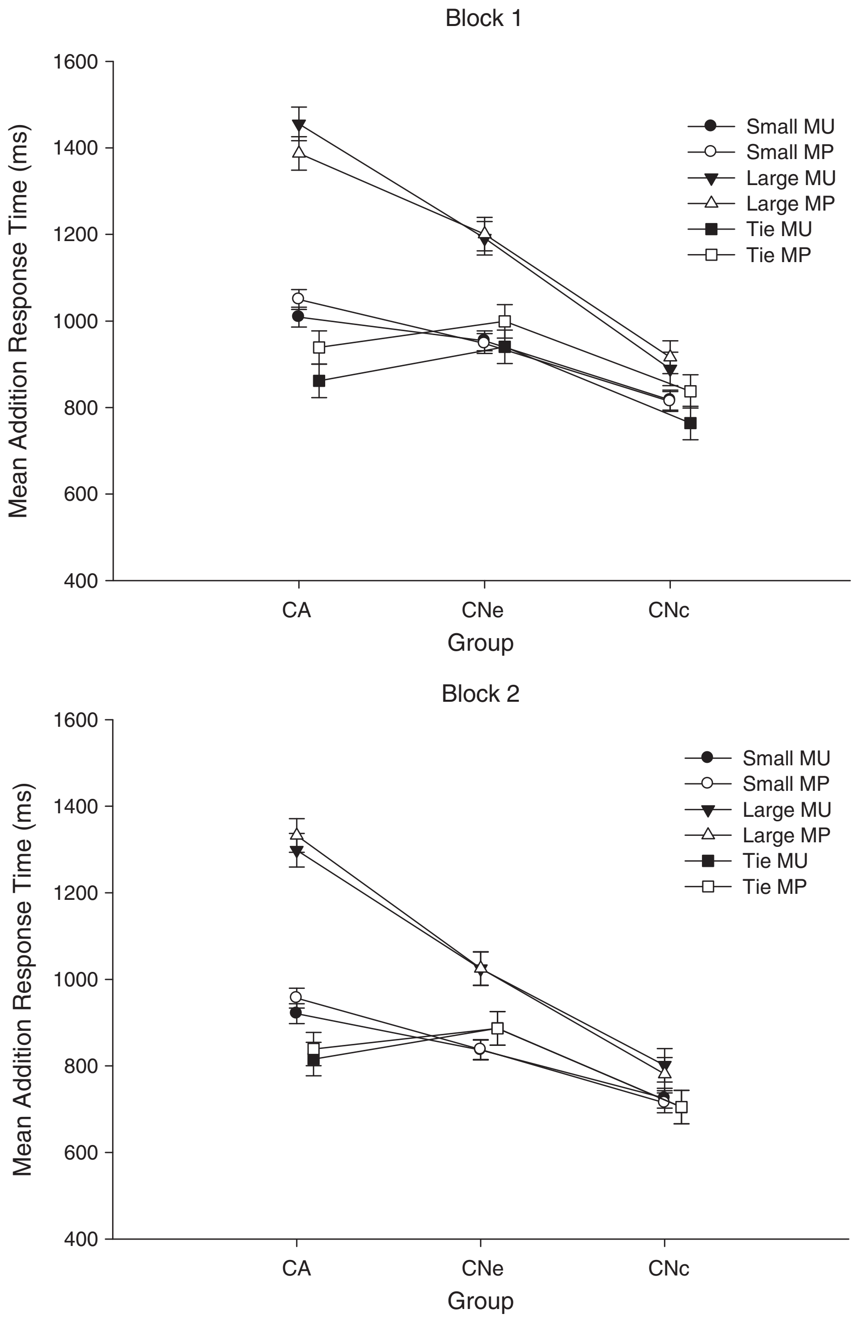

Figure 2. Mean addition response time. Error bars are the 95% repeated-measures confidence intervals based on separate ANOVAs for ties, small, and large problems (Jarmasz & Hollands, 2009; Masson & Loftus, 2003).

Large non-tie addition. CNc answered large non-tie problems fastest (847 ms) followed by CNe (1110 ms) and CA (1368 ms) ($F(2, 105) = 23.8$, $p < .001$, $MSE = 410{,}638$, $\eta_p^2 = .31$), but this effect was qualified by the Group × Set × Block interaction ($F(2, 105) = 4.0$, $p = .02$, $MSE = 13{,}695$, $\eta_p^2 = .07$). Figure 2 shows that for CA the effect of set changed across blocks, with a +68 ms trend for priming in Block 1 and a −34 ms trend for RIF in Block 2 ($F(1, 35) = 4.0$, $p = .05$, $MSE = 46{,}650$, $\eta_p^2 = .10$ for the Block × Set interaction). There were no effects of set on RT for large addition problems for the Chinese participants (all $p > .18$).

Addition retrieval percentage

The strategy reports potentially provided an additional measure of RIF. That is, if practising a multiplication fact (e.g., 4×5) increases subsequent use of procedural strategies for its addition counterpart (e.g., answer $4 + 5$ by counting rather than direct memory retrieval), this implies that memory access for the addition counterpart was impaired (i.e., RIF). CA reported direct memory retrieval (as opposed to a procedural strategy such as counting) for 63.4% of all trials, compared to 78.1% for CNe and 81.6% for CNc. This replicates previous research indicating greater reliance on procedural strategies for addition by Canadian adults than by Chinese adults (Campbell & Xue, 2001). The group differences primarily reflected reported use of counting strategies (CA 18.2%, CNe 6.5%, CNc 3.2%), whereas overall reported use of transformation strategies was similar (CA 17.4%, CNe 15.3%, CNc 14.3%). The "other" strategy option was selected on less than 1% of trials by each group. Figure 3 presents mean percentage of reported use of direct memory retrieval to answer addition problems for each group for tie problems and small and large non-tie problems. The *MSE* from a Group × Block × Set ANOVA for each problem

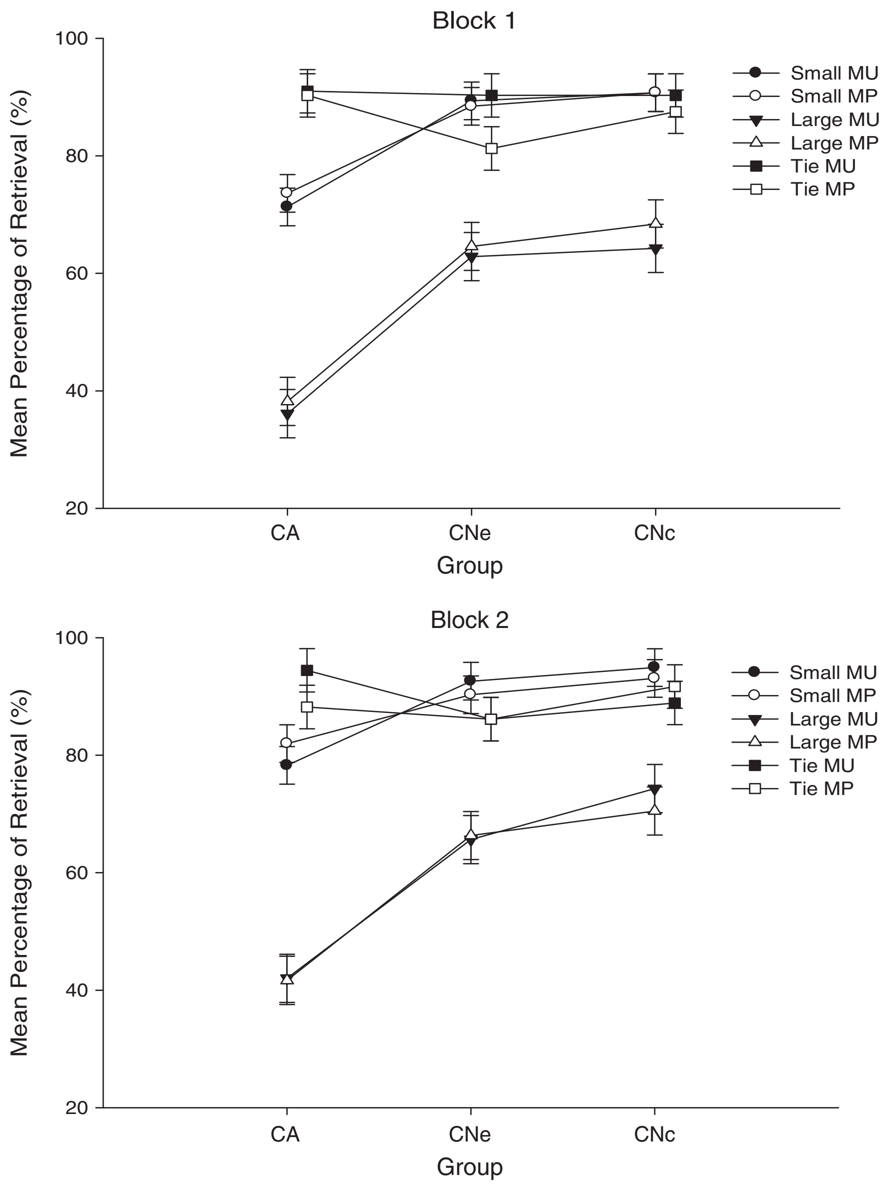

Figure 3. Mean percentage reported retrieval. Error bars are the 95% repeated-measures confidence intervals based on separate ANOVAs for small and large problems (Jarmasz & Hollands, 2009; Masson & Loftus, 2003).

type was used to compute the 95% repeated-measures CIs shown in Figure 3. The *MSE*s for ties, small and large problems were, respectively, 126.1, 98.7, and 150.6.

Tie addition. Direct memory retrieval was reported less often for tie MP addition problems (87.5%) than tie MU addition (90.2%) ($F(1, 105) = 4.7$, $p = .03$, $MSE = 163.0$, $\eta_p^2 = .04$). Thus, there was an overall RIF effect in retrieval reports for tie problems as there was in the RT analysis. The only other significant effect was the Group × Set × Block interaction ($F(2, 105) = 4.1$, $p = .02$, $MSE = 126.1$, $\eta_p^2 = .07$). As Figure 3 shows, this interaction stemmed largely from CNe, who presented a large RIF effect for ties in Block 1 (9.0%, which reflected −4.1% counting and −4.9% transformation for MU ties relative to MP ties).

Small non-tie addition. Whereas the overall rates of reported retrieval for tie problems did not differ across groups, for small non-tie addition, CA reported retrieval less on average (76.3%) than either CNe (90.2%) or CNc (92.4%) ($F(2, 105) = 6.0$, $p = .003$, $MSE = 1830.0$, $\eta_p^2 = .10$). Reported retrieval increased from Block 1 (84.0%) to Block 2 (88.5%) ($F(1, 105) = 12.9$, $p < .001$, $MSE = 167.3$, $\eta_p^2 = .11$). There were no other significant effects (all $p > .19$). Thus, there was no evidence that multiplication practice affected retrieval reports for small non-tie problems.

Large non-tie addition. CA reported far less memory retrieval for large non-tie addition on average (39.5%) than either CNe (64.8%) or CNc (69.4%) ($F(2, 105) = 9.1$, $p < .001$, $MSE = 4102.9$, $\eta_p^2 = .15$; see also Campbell & Xue, 2001). Reported retrieval increased from Block 1 (55.7%) to Block 2 (60.1%) ($F(1, 105) = 6.9$, $p = .01$, $MSE = 134.5$, $\eta_p^2 = .06$). There were no other significant effects (all $p > .10$). Thus, there was no evidence from the ANOVA that multiplication practice affected retrieval reports for the corresponding large, non-tie addition problems.

Addition percentage of errors

The low error rates for tie addition (1.3%) and small non-ties (1.3%) precluded factorial analysis. A Group × Block × Set ANOVA for large addition indicated more errors in Block 1 (3.9%) than in Block 2 (2.4%) ($F(1, 105) = 6.6$, $p = .01$, $MSE = 37.1$, $\eta_p^2 = .06$). Overall, CA tended to make more errors (4.4%) than CNe (2.9%) or CNc (2.3%) ($F(1, 105) = 2.8$, $p = .06$, $MSE = 63.5$, $\eta_p^2 = .05$). There were no other significant effects ($p > .10$).

Relation between response time and percentage retrieval

The 36 RT means in Figure 2 and corresponding retrieval percentages in Figure 3 were highly correlated ($r(34) = -.88$, $p < .001$), indicating a strong negative relationship between mean condition RT and the participants' self-reported use of memory retrieval (see also Campbell & Alberts, 2009). This negative relation is expected because, in general, reported procedure trials are relatively slow compared to retrieval trials (e.g., Campbell & Xue, 2001; LeFevre et al., 1996). Consequently, as percentage of retrieval increases, mean RT decreases. Given this strong relationship, we re-examined RIF effects in RT after excluding reported procedure trials.

Tie addition. For ties, the −39 ms RIF effect in Block 1 (see Figure 2) was eliminated when self-reported procedural trials were excluded (−4 ms; two CA participants were lost because of empty cells). Furthermore, across participants the effect of set (i.e., the MU − MP difference) on tie problems in Block 1 RT and percentage retrieval were negatively correlated ($r(106) = -.21$, $p = .03$). Thus, the expected relation between transfer effects in RT and percentage retrieval held at the level of individual participants in Block 1. In Block 2, this correlation was non-significant ($r(106) = -.05$, $p = .60$). These results suggest that the RIF effect observed in tie RT in Block 1 was associated with strategy-shifts away from retrieval.

Small non-tie addition. For small non-ties the RT analysis with procedure trials excluded reproduced the Group × Set interaction observed in the all trials analysis ($F(2, 99) = 3.9$, $p = .02$, $MSE = 8437$, $\eta_p^2 = .06$). This reflected a −45 ms RIF effect for CA ($F(1, 31) = 5.1$, $p = .03$, $MSE = 8437$, $\eta_p^2 = .14$) but not for either CNc (−1 ms) or the CNe (+16 ms). Four CA and two CNe were lost from this analysis because of empty cells. Unlike the tie problems, for small non-tie problems there was not a significant correlation in Block 1 between effects of set measured in reported retrieval and RT ($r(106) = -.08$, $p = .39$) or in Block 2 ($r(106) = .004$, $p = .97$). This is not surprising given that only CA presented an effect of set (i.e., RIF) on RT for small non-ties.

Large non-tie addition. The corresponding analysis of large addition RTs excluding procedure trials was prohibited by many participants with very few observations per cell and 25 participants lost owing to empty cells. Nonetheless, across participants there was a significant correlation between the effect of set measured by the percentage of reported retrieval and the effect measured in RT both in Block 1 ($r(106) = -.31$, $p = .001$) and Block 2 ($r(106) = -.26$, $p = .006$). The all-trials ANOVA of percentage retrieval for large problems did not indicate an effect of set, but these correlations confirm that the transfer effects (positive or negative) of multiplication practice on RT for large addition problems (i.e., the three-way interaction depicted in Figure 3) were related to corresponding effects on reported retrieval use.

DISCUSSION

To begin, it will be worthwhile to summarise the key transfer findings for each problem type. For tie problems, there was RIF expressed in RT in Block 1 that did not differ among the groups. This effect was largely eliminated when self-reported procedural trials were excluded from the analysis. Furthermore, there was a significant negative correlation between transfer effects (i.e., MU − MP differences) measured in tie-problem RT and percentage retrieval in Block 1 and not in Block 2. Taken together, these results imply that the RIF effect observed in tie RT in Block 1 was due, at least in part, to strategy shifts away from retrieval induced by multiplication practice. Indeed, overall, the percentage of retrieval reported was significantly lower for MP ties than MU ties.

For small non-tie additions, however, only CA presented RIF in RT whereas neither of the Chinese groups did. For the two Chinese groups, a Type (small non-ties vs. ties) × Block × Set analysis of RT confirmed the three-way interaction corresponding to RIF for ties only, and only in Block 1. In contrast, in the corresponding analysis for CA, there was no evidence that RIF differed between ties and small non-ties or that the effect was entirely eliminated in Block 2. This RIF persistence across addition test blocks for the Canadian but not the Chinese participants might indicate that the latter possessed more precise inhibitory control of retrieval competition.

RT for large non-tie problems presented no evidence of RT costs owing to RIF, but there was a Group × Block × Set interaction. This primarily reflected a trend in the CA results for a RT advantage (i.e., priming) of large MP addition relative to MU addition problems in Block 1 that disappeared in Block 2, yielding a significant Set × Block interaction for CA. There was no evidence from the ANOVA that multiplication practice affected retrieval reports for large non-tie problems. Nonetheless, there was a robust negative correlation between the effect of set (i.e., transfer) measured by the percentage of retrieval reported and the effect measured in RT in both addition blocks. This confirmed that the two measures were systematically related in the expected way for large problems.

Procedural hypothesis for addition non-tie problems

We examined potential differences in RIF for tie and non-tie addition problems based on evidence that performance of ties is primarily retrieval-based whereas non-tie addition, including small non-ties such as $2 + 3$ and $3 + 4$, may often be solved by highly automated procedures (Barrouillet & Thevenot, 2013; Fayol & Thevenot, 2012). In this case, we would expect practice of multiplication ties to cause RIF of addition counterparts, whereas practice of multiplication non-ties would yield weaker RIF or fail to produce RIF of their addition counterparts.

The two Chinese groups in the present study produced this pattern of RIF results. Specifically, for the Chinese participants there was a robust RIF effect in RT for addition ties in Block 1 but not for non-ties. The absence of RIF for large non-ties might be attributed to relatively weak memory strength for large additions, as suggested by relatively slow mean RT and high reported use of procedural strategies compared to small non-ties and ties. In this case, large addition non-ties would not strongly compete with their multiplication counterparts during multiplication practice and therefore would not be subject to RIF. The absence of RIF for the small non-ties, however, does not admit to a strength-related explanation because small non-ties and ties were solved with similar efficiency both during multiplication practice and the addition test phase. Self-reported use of procedural strategies for both small non-ties and ties was very low, but Fayol and Thevenot (2012) argued that some addition procedures (e.g., compacted counting procedures, especially those for small addition problems) could be automatised, thereby requiring little conscious attention. Such automated procedures would not be available to consciousness and therefore would not be reflected in participants' strategy self-reports.

In contrast to the Chinese participants, the Canadian group presented RIF for both ties and small non-ties. What could account for this difference in RIF effects between the Canadian and Chinese participants? Language-related differences can be ruled out because, as Campbell and Dowd (2012) observed, both the Chinese group that responded in English and the group the responded in Chinese produced the same RIF pattern. In terms of the compacted procedure theory (Barrouillet & Thevenot, 2013; Fayol & Thevenot, 2012), the differences could indicate that Canadians solved both ties and small non-ties predominantly by direct fact retrieval (making both problem types susceptible to RIF), whereas the Chinese solved ties by direct retrieval but often solved the small non-ties with a procedure (making ties, but not non-ties, susceptible to RIF). We can only speculate about the origins of such a dissociation because there are many culture-related factors that distinguish the development of numerical and arithmetical skills in Canada and China (e.g., see Miller, Kelly, & Zhou, 2005). A potentially unique source of counting-related differences is abacus experience. The abacus is a mechanical counting device (beads on a series of rods are moved using the thumb and fingers) that can be used to perform addition, subtraction, multiplication, division and other arithmetical operations. In China it is popular for children to learn to use an abacus very early in their education (Wang, Geng, Hu, Du, & Chen, 2013). They are encouraged to develop mental abacus skills whereby the states of the abacus are imagined throughout a mental calculation. Regular use of a mental abacus could proceduralise this counting process and ingrain counting into elementary cognitive addition. This process would emerge selectively for addition non-ties but not for addition ties, if the latter are learned as a distinct and salient subcategory of problems, readily available by direct retrieval (Barrouillet & Thevenot, 2013).

Of course, factors other than abacus training could contribute to the group differences in RIF for non-tie additions in the present study. Furthermore, the results do not provide direct evidence for procedure-based performance of the non-ties by the Chinese groups. Nonetheless, the results converge with the operation-sign priming effect for addition non-tie problems, but not tie additions, reported by Fayol and Thevenot (2012), and the larger problem-size effect for small non-tie additions compared to ties reported by Barrouillet and Thevenot (2013). All of these results are consistent with procedural processes for small non-tie additions but not tie additions. Fayol and Thevenot (2012) proposed that development of procedural skills for simple addition (e.g., compacted counting) reflected basic principles of skill acquisition whereby extensive repetition of a task leads from a consciously controlled sequence of steps to a compiled algorithm that is executed automatically when the triggering conditions are encountered. This suggests that such procedures may be linked with level of skill, and Fayol and Thevenot (2012) emphasised that their samples were highly skilled in arithmetic. In this case, the much higher level of arithmetic fluency reflected in French-Kit performance of our Chinese groups compared to the Canadian group could also contribute to differences in procedural skills and the RIF differences observed.

Comparison to previous addition RIF results

The RIF results for the Canadian group are consistent with the corresponding conditions in the two RIF experiments reported by Campbell and Thompson (2012b). These experiments were similar in design to the current study, similarly tested non-Asian Canadians, and used much the same stimuli. For purposes of comparison we pooled the Block 1 data across the two Campbell and Thompson (2012b) experiments including the 42 participants in the retrieval-practice groups (e.g., answered $3 \times 4 = ?$ in the practice phase) and excluding the study-practice groups (e.g., read $3 \times 4 = 12$ silently and then state 12) for whom no RIF was expected or observed.[4] The results indicated RIF in RT for small non-ties (-90 ms) and ties (-57 ms) ($F(1, 41) = 19.1$, $p < .001$, $MSE = 12{,}182$, $\eta_p^2 = .31$) and no evidence that they differed ($F < 1$). The large non-ties presented a trend in the direction of priming rather than RIF ($+69$ ms) ($F(1, 41) = 2.9$, $p = .09$,

[4] We did not include the second addition test block from Campbell and Thompson (2012b) because a block of multiplication intervened between the addition blocks.

MSE =35,073, $\eta_p^2 = .07$). Finally, there was no clear evidence for an effect of set (i.e., RIF or priming) on retrieval reports or error rates. Thus, the RIF ANOVA results of the present Canadian group replicated the RIF pattern observed by Campbell and Thompson (2012b) with their non-Asian Canadian groups.

With respect to RIF results for the Chinese groups, unlike the present study, Campbell and Dowd's (2012) experiment with Chinese-English bilinguals presented evidence of RIF in RTs for both small and large problems, once the small rule-based problems were excluded (Campbell & Therriault, 2013). This effect occurred only when practice language and test language were the same, as they were in the present study. A straightforward interpretation of the RIF differences between these experiments, however, is precluded by methodological differences. Unlike the present study, there was an addition pre-test of both MP and MU addition problems performed in either Chinese or English, and participants practised two different multiplication problem sets, one in Chinese and one in English. That experiment utilised three problem sets that included only one or two tie problems, ruling out a meaningful analysis of ties only. An analysis of small non-ties similarly has low power and precision once the ties and rule-based problems are excluded (there are only three observations per cell). With respect to large problems, the Chinese participants in the present study were 28% slower during the first addition test block (1049 ms for the Chinese groups test here compared to 822 ms in Campbell and Dowd, 2012). This could reflect methodological or sampling differences, but in either case it raises the possibility that RIF for large problems occurred in Campbell and Dowd's (2012) study because their participants had exceptionally well-developed memory for the large addition facts, which attracted inhibition and RIF during multiplication practice.

Canadian vs. Chinese simple arithmetic performance

With respect to arithmetic performance, the results replicated previous research comparing Canadian and Chinese participants (e.g., Campbell & Xue, 2001; LeFevre & Liu, 1997). The Chinese participants were faster overall but especially so for large problems (Figures 1 and 2). For addition, the Chinese groups reported substantially greater use of direct memory retrieval than the Canadians, and this difference was greatest for non-tie problems, especially large non-ties (see also Campbell & Gunter, 2002). The Chinese groups' performance was similar for ties and small non-ties whereas the Canadians were faster and more likely to report retrieval for ties than non-ties, a pattern also observed by Campbell & Gunter (2002). As retrieval is often faster than self-reported procedural strategies, greater use of retrieval for ties would contribute to their faster addition RTs (Campbell & Xue, 2001). Although self-reported procedures are often slow compared to direct retrieval, Barrouillet and Thevenot (2013, p. 43) noted that automated procedures, such as compacted counting, might be executed very quickly. Consequently, the faster RT observed for small non-ties by the Chinese groups compared to the Canadian group (Figure 2) is not inherently at odds with the possibility that Chinese performance of these problems was based on procedures whereas Canadian performance was based on fact retrieval.

Finally, as Campbell and Dowd (2012) observed, Chinese participants who responded in English were slower compared to those responding in Chinese, and this difference was exaggerated for large addition problems (see also Campbell & Epp, 2004). This could reflect less efficient direct access to large problems in English or that translation from Chinese to English number words is slower for large sums (i.e., 11 20), which are idiosyncratic in English (eleven, twelve, thirteen, etc.) but regularised in Chinese (ten-one, ten-two, ten-three, etc.).

CONCLUSIONS

The present study replicated previously reported differences in the arithmetic performance of Canadian and Chinese adults (e.g., Campbell & Gunter, 2002; Campbell & Xue, 2001; LeFevre & Liu, 1997), indicating that these are reliable and persistent phenomena. The analysis of RIF uncovered additional subtle group differences in the RIF effects induced by practising multiplication counterparts of the addition test problems. Tie problems produced a RIF effect in the first addition block (i.e., slower RT for MP compared to MU additions) that did not differ across groups, but only the Canadian group showed RIF for the small non-tie problems. As procedure-based arithmetic does not foster problem-specific RIF (Campbell & Therriault, 2013), the Chinese groups' RIF effect for ties but not small non-ties is consistent with the use of procedural skills for small non-tie additions. The results converge with other evidence (Barrouillet & Thevenot 2013; Fayol & Thevenot, 2012) that addition ties are solved by direct memory retrieval, making them subject to RIF, whereas small non-ties may be solved by highly efficient procedural processes, particularly in highly skilled individuals.

Original manuscript received October 2012
Revised manuscript received May 2013
First published online August 2013

REFERENCES

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language*, *49*, 415–445. doi:10.1016/j.jml.2003.08.006

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1063–1087. doi:10.1037/0278-7393.20.5.1063

Barrouillet, P., & Thevenot, C. (2013). On the problem-size effect in small additions: Can we really discard any counting-based account? *Cognition*, *128*, 35–44. doi:10.1016/j.cognition.2013.02.018

Campbell, J. I. D., & Alberts, N. A. (2009). Operation-specific effects of numerical surface form on elementary calculation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 999–1011. doi:10.1037/a0015829

Campbell, J. I. D., & Dowd, R. R. (2012). Inter-operation transfer in Chinese-English bilinguals' arithmetic. *Psychonomic Bulletin & Review*, *19*, 948–954. doi:10.3758/s13423-012-0277-z

Campbell, J. I. D., & Epp, L. J. (2004). An encoding-complex approach to numerical cognition in Chinese-English bilinguals. *Canadian Journal of Experimental Psychology*, *58*, 229–244. doi:10.1016/S0010-0277(98)00006-7

Campbell, J. I. D., & Gunter, R. (2002). Calculation, culture, and the repeated operand effect. *Cognition*, *86*, 71–96. doi:10.1016/S0010-0277(02)00138-5

Campbell, J. I. D., & Metcalfe, A. W. S. (2007). Numeral format and arithmetic rules. *European Journal of Cognitive Psychology*, *19*, 335–355. doi:10.1080/09541440600717610

Campbell, J. I. D., & Phenix, T. L. (2009). Target strength and retrieval-induced forgetting in semantic recall. *Memory & Cognition*, *37*, 65–72. doi:10.3758/MC.37.1.65

Campbell, J. I. D., & Therriault, N. H. (2013). Retrieval-induced forgetting of arithmetic facts but not rules. *Journal of Cognitive Psychology*. doi:10.1080/20445911.2013.798328

Campbell, J. I. D., & Thompson, V. A. T. (2012a). *MorePower 6.0* for ANOVA with relational confidence intervals and Bayesian analysis. *Behavior Research Methods*, *44*, 1255–1265. doi:10.3758/s13428-012-0186-0

Campbell, J. I. D., & Thompson, V. A. T. (2012b). Retrieval-induced forgetting of arithmetic facts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 118–129. doi:10.1037/a0025056

Campbell, J. I. D., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General*, *130*, 299–315. doi:10.1037/0096-3445.130.2.299

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

Fayol, M., & Thevenot, C. (2012). The use of procedural knowledge in simple addition and subtraction problems. *Cognition*, *123*, 392–403. doi:10.1016/j.cognition.2012.02.008

Galfano, G., Penolazzi, B., Fardo, F., Dhooge, E., Angrilli, A., & Umiltà, C. (2011). Neurophysiological markers of retrieval-induced forgetting in multiplication fact retrieval. *Psychophysiology*, *48*, 1681–1691. doi:10.1111/j.1469-8986.2011.01267.x

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 & Review*, *4*, 425–430. doi:10.3758/BF03210805

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. doi:10.1037/0012-1649.32.2.254

Grabner, R., & De Smedt, B. (2011). Neurophysiological evidence for the validity of verbal strategy reports in mental arithmetic. *Biological Psychology*, *87*, 128–136. doi:10.1016/j.biopsycho.2011.02.019

Imbo, I., & LeFevre, J.-A. (2010). The role of phonological and visual working memory in complex arithmetic for Chinese- and Canadian-educated adults. *Memory & Cognition*, *38*, 176–185. doi:10.3758/MC.38.2.176

Jakab, E., & Raaijmakers, J. G. W. (2009). The role of item strength in retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 607–617. doi:10.1037/a0015264.

Jarmasz, J., & Hollands, J. G. (2009). Confidence intervals in repeated-measures designs: The number of observations principle. *Canadian Journal of Experimental Psychology*, *63*, 124–138. doi:10.1037/a0014164

Jost, K., Beinhoff, U., Hennighausen, E., & Rösler, F. (2004). Facts, rules, and strategies in single-digit multiplication: Evidence from event-related brain potentials. *Cognitive Brain Research*, *20*, 183–193. doi:10.1016/j.cogbrainres.2004.02.005

Kirk, E. P., & Ashcraft, M. H. (2001). Telling stories: The perils and promise of using verbal reports to study math strategies. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, *27*, 157–175. doi:10.1037/0278-7393.27.1.157

LeFevre, J.-A., & Liu, J. (1997). The role of experience in numerical skill: Multiplication performance in adults from China and Canada. *Mathematical Cognition*, *3*, 31–62. doi:10.1080/135467997387470

LeFevre, J.-A., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 216–230. doi:10.1037/0278-7393.22.1.216

Levy, B. J., McVeigh, N. D., Marful, A., & Anderson, M. C. (2007). Inhibiting your native language: The role of retrieval-induced forgetting during second-language acquisition. *Psychological Science*, *18*, 29–34. doi:10.1111/j.1467-9280.2007.01844.x

Maslany, A. J., & Campbell, J. I. D. (2013). Failures to replicate hyper-retrieval-induced forgetting in cognitive arithmetic. *Canadian Journal of Experimental Psychology*, *67*, 73–77. doi:10.1037/a0031138

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology*, *57*, 203–220. doi:10.1037/h0087426

Metcalfe, A. W. S., & Campbell, J. I. D. (2011). Strategies for simple addition and multiplication: Verbal self-reports and the operand recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 661–672. doi:10.1037/a0022218

Miller, J. (1988). A warning about median reaction time. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 539–543. doi:10.1037/0096-1523.14.3.539

Miller, K. F., Kelly, M., & Zhou, X. (2005). Learning mathematics in China and the United States: Cross-cultural insights into the nature and course of preschool mathematical development. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 163–178). New York, NY: Psychology Press.

Norman, K. A., Newman, E. L., & Detre, G. (2007). A neural network model of retrieval-induced forgetting. *Psychological Review*, *114*, 887–953. doi:10.1037/0033-295X.114.4.887

Pesenti, M., Depoorter, N., & Seron, X. (2000). Noncommutability of the $N + 0$ arithmetical rule: A case study of dissociated impairment. *Cortex*, *36*, 445–454. doi:10.1016/S0010-9452(08)70853-0,

Phenix, T. L., & Campbell, J. I. D. (2004). Effects of multiplication practice on product verification: Integrated structures model or retrieval-induced forgetting? *Memory & Cognition*, *32*, 324–335. doi:10.3758/BF03196862

Seyler, D. J., Kirk, E. P., & Ashcraft, M. H. (2003). Elementary subtraction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 1339–1352. doi:10.1037/0278-7393.29.6.1339

Storm, B., & Levy, B. (2012). A progress report on the inhibition account of retrieval-induced forgetting. *Memory & Cognition*, *40*, 827–843. doi:10.3758/s13421-012-0211-7

Thevenot, C., Castel, C., Fanget, M., & Fayol, M. (2010). Mental subtraction in high and lower-skilled arithmetic problem solvers: Verbal report vs. operand-recognition paradigms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1242–1255. doi:10.1037/a0020447

Wang, Y., Geng, F., Hu, Y., Du, F., & Chen, F. (2013). Numerical processing efficiency improved in experienced mental abacus children. *Cognition*, *127*, 149–158. doi:10.1016/j.cognition.2012.12.004

Zbrodoff, N. J., & Logan, G. D. (2005). What everyone finds: The problem size effect. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 331–346). New York, NY: Psychology Press.