Journal of Experimental Child Psychology 122 (2014) 62–74

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Specificity of children's arithmetic learning

Drew Walker, Daniel Bajic, Laura Mickes, Jung Kwak, Timothy C. Rickard *

Department of Psychology, University of California, San Diego, La Jolla, CA 92093, USA

ARTICLE INFO

Article history:
Received 26 September 2012
Revised 4 November 2013
Available online 8 February 2014

Keywords:
Memory
Learning
Transfer
Identical elements model
Arithmetic
Problem solving

ABSTRACT

Among adults, arithmetic training–transfer studies have documented a high degree of learning specificity. Provided that there is a delay of at least 1 day between training and testing, performance gains do not transfer to untrained problems, nor do they transfer to complement operation-inverted problems (e.g., gains for 4 + 7 = __ do not transfer to the complement subtraction problem, 11 − 4 = __, or vice versa). Here we demonstrate the same degree of learning specificity among 6- to 11-year-old children. These results appear to rule out, for the current training paradigm, operation-level procedural learning as well as any variant of complement problem mediation that would predict transfer. Results are consistent with either or both of two types of learning: (a) item-level procedural learning and (b) a shift to memory-based performance as predicted by the elemental elements model. These results suggest a developmental pattern such that specificity of learning among children is similar to that among adults. Educational implications are noted.

Introduction

Answer production practice, or drill, clearly improves adults' and children's arithmetic performance (e.g., Burns, 2005; Fendrich, Healy, & Bourne, 1993). Among adults, that improvement is largely specific to trained problems. Surprisingly, however, the specificity of children's arithmetic fluency training has not been explored in the literature. For example, there appears to be no experimental work among children that addresses the extent to which extended training on a subset of multiplication problems (e.g., 4 × 7) transfers to complementary division problems (e.g., 28 ÷ 4) or vice versa, nor has the analogous

* Corresponding author.
E-mail address: trickard@ucsd.edu (T.C. Rickard).

http://dx.doi.org/10.1016/j.jecp.2013.11.018
0022-0965/

work been done for addition and subtraction (but see Baroody, 2010, for exploration of conceptual specificity for addition and subtraction among young children). Given the substantial student time investment in skill training, it is important to document this transfer (or specificity) of learning phenomena.

Summary of adult findings

Among adults, arithmetic learning in delayed training–transfer studies (i.e., studies in which at least 1 day intervenes between training and the transfer test) is highly specific to trained problems. Bajic, Kwak, and Rickard (2011) showed that training gains for a subset of addition and subtraction problems do not transfer to untrained problems, nor is there cross-operation transfer to complement problems; that is, training gains for 4 + 7 = __ do not transfer to the complement subtraction problem, 11 – 4 = __, or vice versa. The analogous specificity pattern holds for multiplication and division (Rickard & Bourne, 1996).

In immediate transfer paradigms, however, there is often evidence of partial cross-operation transfer among adults (cf. Campbell, Fuchs-Lacelle, & Phenix, 2006). Campbell and Agnew (2009) observed addition–subtraction transfer on a test that immediately followed repetition training. Cross-operation transfer for multiplication–division has also been demonstrated in immediate (or very brief delay) priming experiments (Campbell, 1999; LeFevre & Morris, 1999). Thus, adult arithmetic can be characterized as exhibiting nearly complete specificity of learning on delayed transfer tests and partial transfer of learning on immediate tests. Bajic et al. (2011) proposed that partial transfer on immediate transfer tests may reflect the influence of episodic memories of trained problems that are relatively quickly forgotten. On delayed transfer tests, performance may be driven by better retained but also more specific learning that does not support cross-operation transfer. From educational and developmental points of view, these better retained representations are arguably of greater interest and they are the focus of this study.

Overview of the experimental design

To investigate delayed transfer effects among children, 6- to 11-year-old students were trained for six sessions on either a set of mixed addition and subtraction problems or a set of mixed multiplication and division problems, depending on their pre-experimental skill level. On identical pre- and posttests performed in separate sessions, performance was evaluated on each of three problem sets: (a) *trained* problems, (b) *operation-inverted complements* of trained problems (e.g., trained on 4 × 7 and tested on 28 ÷ 4 or vice versa), and (c) *untrained* problems (i.e., problems not trained in either operation, e.g., 3 × 8). Two main empirical questions were addressed. First, do performance improvements on trained problems transfer to untrained problems from the same operation? Second, compared with untrained problems, is there greater transfer of learning to operand-inverted problems?

Candidate arithmetic learning mechanisms and transfer predictions

The literature on adults' and children's arithmetic suggests several mechanisms that may be involved in performance and learning, with varying predictions for transfer. These mechanisms are grouped into four categories in Table 1. The first two categories involve procedural learning. In both of those categories, learning can, in principle, occur either as a *shift* from a less efficient procedure to a more efficient procedure or as *improved efficiency* in procedural step execution (e.g., faster counting for addition). Multiple types of procedural shifts have been identified (e.g., Robinson et al., 2006; Siegler & Jenkins, 1989). For addition, children may use finger counting initially, then progress to counting on from one of the addends, and then progress to the more efficient strategy of counting on from the larger addend (the *min rule*; Groen & Parkman, 1972). For multiplication, the repeated addition strategy exhibits a developmental shift toward adding the larger operand to itself the number of times indicated by the smaller one (Lemaire & Siegler, 1995).

Operation-level procedural learning refers to procedural learning that will transfer from trained problems to other problems from the same operation, but not to problems from any of the other operations. Intuitively at least, most procedural learning would be expected to be operation-level learning. Consider procedural shifts. Once a student shifted to the min rule for a given addition problem, one

Table 1
Candidate learning mechanisms and corresponding transfer predictions for the current experiment.

Learning mechanism	Transfer to operation inversion?	Transfer to untrained?
Operation-level procedures[a,b]	Yes	Yes
Item-level procedures[a]	No	No
Complement problem mediation[c]	Yes	No
Subtraction by addition		
Division by multiplication		
Division by factoring		
Shift to, or improved execution of, identical elements retrieval	No	No

[a] Procedural learning in the form of either procedural shifts or improved execution.
[b] Although operation-level procedures as defined in this article support transfer only from one subset of problems to another subset from the same operation, transfer to the inverted operation is nevertheless predicted in the current experiment because training occurs on two operations and problems on the transfer test are drawn exclusively from the same two operations.
[c] These predictions apply only to the variant of the complement problem mediation hypothesis that is described in the text.

might expect rapid generalization of that rule to other addition problems. Similarly, a shift to repeatedly adding the larger operand rather than the smaller operand in multiplication would be expected to generalize to untrained multiplication problems. Improved efficiency of step execution for a given procedure might also be expected to generalize from trained problems to untrained problems from the same operation (e.g., faster repeated addition for 6×7 would be expected to transfer to 4×7).

Given the design of the current experiment, in which each student was trained on a subset of both addition and subtraction (or multiplication and division) problems, the operation-level procedural learning account predicts transfer of learning to both the untrained problems and operation-inverted conditions (because both of those conditions exclusively contain problems from the two operations on which the students were trained).

Item-level procedural learning, which as defined here would not transfer beyond a particular arithmetic problem, may also occur. For example, generalization of the addition min rule from one problem to another can be delayed (Shrager & Siegler, 1998), indicating that students sometimes discover the min rule independently for a number of problems before they generalize it as an operation-level procedure. Hence, temporarily at least, discovery of the min rule may take the form of item-level learning. It is also possible that some aspects of procedural learning in the form of improved efficiency, such as improved repeated addition for a particular multiplication problem, do not generalize to other problems. Item-level procedural learning predicts, by definition, no transfer in the current experiment to either operation-inverted or untrained problems.

A third category is referred to here as *complement problem mediation,* which involves solving a problem by accessing knowledge of its inverse operation complement. Complement problem access may, in principle, occur through use of either procedures or memory retrieval. Three types of complement problem mediation have been hypothesized: subtraction by addition (Campbell, 2008; Peters, De Smedt, Torbeyns, Ghesquire, & Verschaffel, 2010), division by multiplication (Campbell, 1999; Mauro, LeFevre, & Morris, 2003), and division by factoring (Campbell & Robert, 2008; Rickard, 2005; Rusconi, Galfano, Rebonato, & Umiltà, 2006). All types are founded on the assumptions that (a) knowledge about addition (or multiplication) typically exceeds that about the inverse operation (subtraction or division), and (b) under at least some circumstances, subtraction and division performance for a given problem can be facilitated (relative to use of other strategies) by accessing knowledge about the inverse complement. Evidence consistent with complement problem mediation comes from the immediate transfer and priming experiments discussed earlier, verbal reports that imply mediation (Barrouillet, Mignon, & Thevenot, 2008; Robinson, 2001; Robinson et al., 2006), and problem format effects among adults (Campbell, 2008; Campbell & Alberts, 2010; Mauro et al., 2003).

The concept of complement problem mediation has not yet been defined sufficiently to allow for unqualified transfer of learning predictions in the current experiment. As often treated in the literature, however, positive cross-operation transfer is implied (e.g., Campbell, 2008). The assumption in that variant of the mediation hypothesis is that a common problem representation is accessed and strengthened when performing either complement problem. For example, training on the addition problem 4 + 7 = __ is assumed to strengthen a flexibly accessible addition representation, *4 + 7 = 11*. That strengthening is assumed to facilitate processing through that representation for any presented problem that can access it, such as a complement problem in the inverted operation. Thus, when the problem 11 − 4 = __ is presented on the transfer test, complement problem mediation via *4 + 7 = 11* is facilitated by the prior addition training, yielding cross-operation transfer from addition to subtraction. Similarly, on each trial on which complement problem mediation is used for subtraction during training, the underlying common addition representation (e.g., *4 + 7 = 11*) is strengthened, supporting cross-operation transfer to 4 + 7 = __. The transfer prediction for complement problem mediation in Table 1 refers to this particular variant and is not intended as a global prediction applicable to other possible variants.

The fact that cross-operation transfer has not been observed in delayed transfer experiments among adults suggests that complement problem mediation either (a) does not occur for adults in training paradigms or (b) occurs but for some reason does not yield cross-operation transfer (i.e., the variant of the mediation concept described above is not correct). The current study provides insight into whether the same conclusions hold for children.

Finally, Rickard, Healy, and Bourne (1994) and Rickard and Bourne (1996) proposed what they termed identical elements (IE) representations that preclude memory-based retrieval through complement problem representations. Their model proposes that practice yields a separate and independent memory representation for each unique combination of practiced stimulus elements. Each stimulus representation, in turn, has an association with the required response (e.g., 4, 7, × → 28; 28, 7, ÷ → 4). Critically, an IE representation can be accessed only by items that have an exactly matching set of stimulus elements, including the logical arithmetic operation required but ignoring superficial perceptual factors such as stimulus modality, the operation symbol, and (for commutative operations) the spatial or temporal ordering of the elements (Rickard & Bourne, 1996). Hence, shifts to or strengthening of IE memory would produce no transfer to either operation-inverted or untrained problems.

Specificity of learning across lifespan development

In addition to their implications for cognitive process accounts, the results of the current transfer experiment should yield insight into general principles of learning specificity across development. The specificity of learning among children may plausibly be related to that among adults in one of three ways. First, substantial transfer to inverted problems, and perhaps also to untrained problems, may be observed. For example, children may undergo operation-level strategy shifts or make frequent use of complement problem mediation that supports transfer, whereas new learning among adults may be more item based (e.g., may primarily reflect IE-based memory retrieval), leading to greater specificity. Second, learning specificity may be a developmental constant, such that children's specificity patterns closely match those of adults, raising the possibility that, during independent training, the strategies used and the corresponding learning are analogous for the two populations. A third possibility, for completeness, is that children may have an even higher level of learning specificity than do adults. Children are likely to have encountered arithmetic and other skills in less varied contexts than have adults. As a result, children's mental representations may be more tied to specifics such as problem format and arithmetic operation than are those for adults. Given that adults exhibit no transfer in the delayed transfer paradigm explored here, the current results are not expected to differentiate between this possibility and the developmental constant account.

Method

Setting and participants

The experiment was conducted as part of an after-school Math Club program held twice a week for six consecutive weeks at a community center in southeastern San Diego County on the U.S. West

Coast. All students provided voluntary written assent to participate, and all parents provided voluntary written consent. The Math Club recruits students from all schools within a school district where 54% of students are enrolled in the free or reduced lunch program and where 59% of students are categorized at the basic or above level of mathematics. The ethnicity distribution of that school district is 44.7% Hispanic or Latino, 32.5% White not Hispanic, 10.6% African American, 2.3% Asian, 5.8% two or more races, 2.8% Filipino, and 0.4% American Indian or Alaska Native.

A sample of 38 students aged 6 to 11 years attended at least one training session and the posttest. Two students who had posttest accuracy of less than 50% were excluded. Exclusion of those students did not alter the statistical results. Of the remaining 36 students, 22 had signed up for Math Club early enough to also attend the pre-assessment and pretest sessions. Most students who did not attend the pretest were in the 6- to 8-year-old range. That pattern reflects the fact that Math Club was advertised for the younger ages at a later time than for the older ages, a factor that was not under our control.

Design and materials

There were nine experimental sessions: one on Thursday of the first week of Math Club and one session each on Wednesday and Thursday of the next 4 weeks. Each session lasted approximately 30 min. The pre-assessment took place during Session 1, the pretest during Session 2, training during Sessions 3 to 8, and the posttest during Session 9.

Stimuli for each student were generated from one of four master stimulus sets, each of which was composed of 24 arithmetic number triplets. As shown in the Appendix, the four master sets were classified as "easy" addition–subtraction, "difficult" addition–subtraction, "easy" multiplication–division, and "difficult" multiplication–division. Each master set was divided into three subsets (A, B, and C), such that the eight number triplets in each subset were roughly balanced on sum or product size.

During training, each student was presented repeatedly with 16 arithmetic problems taken from two of the three subsets of the student's assigned master set (i.e., Subsets A and B, Subsets B and C, or Subsets C and A, counterbalanced over students). For each student, the eight triplets from one of the subsets were presented throughout training as addition (or multiplication) problems and the eight triplets from the other subset were presented in the inverse operation (subtraction or division), with counterbalancing over participants. Half of the problems from each subset were randomly selected for presentation with ascending operand order throughout training, and half were presented with descending order (e.g., the addition problem 2 + 4 = __ has ascending operand order).

The pretest and posttest materials consisted of 48 problems taken from the student's master set (one for each operation from each of the 24 triplets; see Appendix). As was the case during training, half of the problems were presented in ascending operand order and half in descending order (for the 16 trained problems, the presented operand order matched that during training). This design yielded "pre" and "post" data for each of three transfer conditions: (a) the 16 *trained* problems, (b) the 16 *operation inversions* of the trained problems, and (c) the 16 *untrained* problems taken from the subset of eight number triplets that were not presented during training.

All phases of the experiment were performed using paper and pencil. Each student received a binder containing a set of worksheets on which problems were presented in horizontal fill-in-the-blank format (e.g., $4 \times 7 =$ __, $28 \div 4 =$ __). On each worksheet, 16 problems (eight from each complement operation) were printed in two columns of eight. Problem ordering on each sheet was randomly determined. Problems were printed only on the front side of each sheet.

Procedure

In Session 1 (pre-assessment), students worked through as many problems as they could in each of four 2-min performance blocks: one block each for problems created from the easy addition–subtraction, difficult addition–subtraction, easy multiplication–division, and difficult multiplication–division master sets. In each block, the 48 problems (24 from each operation) from the master set were randomly distributed over four pages of 16 problems, with the constraint that there were 8 problems from each operation on each page. This pre-assessment allowed us to identify and assign to each student the most difficult master set that the student could perform while exhibiting high accuracy

(>70%) and no systematic errors that would reflect a lack of conceptual understanding (e.g., performing addition on multiplication problems or consistently making the same errors on a subset of problems within the master set). Students who enrolled in Math Club after the pre-assessment session were assigned to a master set that, based on grade-level curriculum expectations, they should be able to perform accurately with no systematic error (i.e., 6- and 7-year-olds: easy addition–subtraction; 8-year-olds: difficult addition–subtraction; 9-year-olds: easy multiplication–division; 10- and 11-year-olds: difficult multiplication–division).

From the pretest session onward, students worked problems in 3-min blocks with 2-min breaks between blocks. Research assistants ensured that students worked continuously throughout each block. During the pretest and posttest sessions, there were three blocks, each corresponding to a different transfer condition: one block in which only the 16 trained problems were presented, one block in which only the 16 operation-inverted problems were presented, and one block in which only the 16 untrained problems were presented. Block order was counterbalanced over students. During training, there were six sessions and six blocks per session. On all training blocks, participants exclusively worked the 16 problems from their training set.

On each block, students were instructed to complete as many worksheets as possible until told to stop. Each worksheet contained one instance of each of the 16 problems to be presented during the block, randomly ordered anew on each worksheet. More worksheets were included for each block than students were able to complete. That feature of the experiment served two purposes. First, it allowed the number of problems completed during each 3-min block to be used as a measure of performance speed. Second, because all students worked continuously throughout each training period, students who worked more slowly were not exposed to direct performance evaluation from students who worked more quickly.

To begin each block, the experimenter instructed students to turn their binder to a separation page that indicated the beginning of a block of problems. The experimenter held up a stopwatch, asked for silence, and then instructed the students to begin, at which point students flipped to their first worksheet page and began working. At the end of 3 min, the experimenter instructed the students to stop. A team of assistants made sure that all of the students put their pencils down at that point. Likewise, by the end of each break period, the assistants made certain that all students had found the separation page that marked the start of the next block.

Data analysis

The dependent variables for each training and test block were *number correct* (total number of problems correctly solved during the allotted time for each block) and *accuracy* (proportion of problems attempted during each block that were correctly solved). Number correct is likely to be the more sensitive of these dependent measures. Statistical analyses were also performed on accuracy, however, to provide insight into whether the transfer findings manifest in both the rate of problem solving (i.e., number correct) and accuracy.

Posttest analyses on the full set of 36 students involved factors of age (between participants), transfer condition (within participants), and their interaction. Pre–Post analyses involving the factors of test (pre vs. post) and transfer condition were also performed on the subset of 22 students who had also attended the pretest session. Prior results with samples of 24 adults (e.g., Bajic et al., 2011; Rickard & Bourne, 1996) suggest that, with $N = 36$ for the posttest in the current study, statistical power to detect the within-participants effects (the main effect of transfer condition and the age by transfer condition interaction) will be high. This extrapolation from adults to children, however, presumes similar within-participants variance patterns for the two populations and, thus, can provide only a rough guide.

We expected the power to detect any between-groups effect of age to be relatively low. However, the study was not intended to address whether there is a main effect of age or of the different problem sets and arithmetic operations on which children of various ages were trained. Rather, the design goals were to determine (a) whether the main effect of transfer condition that has been observed for adults also holds for children and (b) whether the answer to the former question must be qualified by age.

Our use of different problem sets and arithmetic operations across ages was motivated by the design goal of evaluating whether specificity of learning holds across ages for the case in which children of each age are able to perform the problems with high accuracy but, prior to experimental training, are unlikely to have achieved optimal procedural strategy execution or to have shifted to memory-based performance for most problems. It would not be possible to evaluate this issue over ages without allowing the problems sets and operations on which the children were trained to vary.

Results

Among the 36 students who attended at least one training session and the posttest, there were 1, 9, 9, 8, 5, and 4 students of ages 6, 7, 8, 9, 10, and 11 years, respectively. Because preliminary analysis indicated no difference in relative condition performance between 10- and 11-year-old students, those nine students were combined into one 10–11-year-old category for all subsequent analyses.

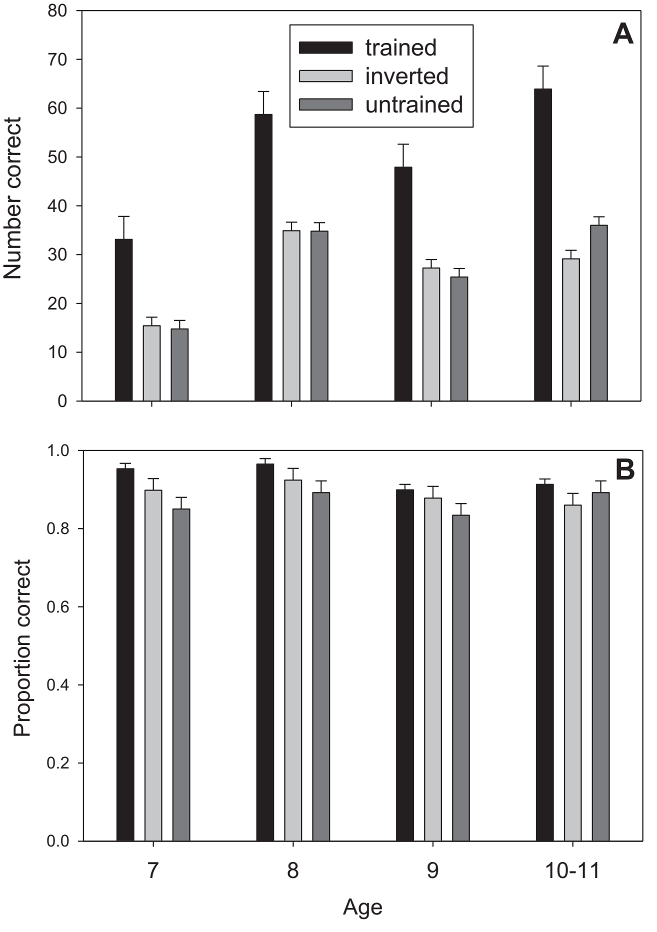

Fig. 1. Posttest results for number correct (A) and accuracy (B) plotted by age and transfer condition. Error bars for the trained problems represent the standard errors based on the within-participants error term of the ANOVA comparing the trained problems condition with the mean of the untrained and operation-inverted problems conditions. Error bars for the inverted and untrained problems represent standard errors based on the within-participants error term of the ANOVA comparing the untrained problems condition with the operation-inverted problems condition.

The single 6-year-old student was also combined with the 7-year-old group. This grouping yielded 10, 9, 8, and 9 students in the 7-, 8-, 9-, and 10–11-year-old groups, respectively. Of this sample, seven students were assigned to Master Set 1, 16 to Master Set 2, 2 to Master Set 3, and 11 to Master Set 4.

Of the 36 students, 22 also attended the pretest session. Among that subgroup, 0, 2, 5, 8, 4 and 3 students were ages 6, 7, 8, 9, 10, and 11 years, respectively. Of these students, 12 were assigned to Master Set 2, 1 to Master Set 3, and nine to Master Set 4.

Training

The 36 students attended an average of 4.8 of the six training sessions (range = 1–6). The mean numbers of problems correctly solved were 29.2, 29.6, 36.6, 40.1, 47.6, and 43.8 in Sessions 1, 2, 3, 4, 5, and 6, respectively. Mean accuracy levels, averaged over the six blocks within each session, were .89, .93, .94, .94, .93, and .94 from the first session to the sixth session. These high accuracies verify that our efforts to assign each student a master set that the student could perform well under independent training circumstances was successful. Similar learning curves were exhibited for each grade level. A one-way within-participants analysis of variance (ANOVA) confirmed a strong effect of session on number correct, $F(5,133) = 17.2$, $p < .0001$, $\eta_p^2 = .39$.

Posttest

Number correct

Posttest results for the full set of 36 students are shown in Fig. 1A as a function of age and transfer condition. Two planned orthogonal factorial ANOVAs, with factors of age (7, 8, 9, or 10–11 years; between participants) and transfer condition (within participants), were performed on number correct.[1] Each ANOVA implements a single degree of freedom contrast involving the transfer factor, efficiently testing the critical predictions of the candidate learning mechanisms described in Table 1. Those contrasts have been shown to capture the great majority of the condition variance in prior studies of delayed transfer among adults (Bajic et al., 2011; Rickard & Bourne, 1996).

In the first ANOVA, the number correct in the trained condition was contrasted against the number correct averaged over the inverted and untrained problems conditions. There was no significant effect of age, $F(3,32) = 1.95$, $p = .14$, $\eta_p^2 = .15$. There was a highly significant effect of transfer condition, $F(1,32) = 29.36$, $p < .0001$, $\eta_p^2 = .48$, but no interaction between age and transfer condition, $F(3,39) < 1.0$. These results confirm the dominant pattern in Fig. 1A; training effects were much greater for trained problems than for inverted and untrained problems.

In the second ANOVA, the number correct in the operation-inverted problems condition was contrasted with the number correct in the untrained problems condition, with the trained problems condition excluded. There was again no significant effect of age, $F(3,32) = 2.17$, $p = .11$, $\eta_p^2 = .17$. There was also no significant effect of either transfer condition, $F(1,32) < 1.0$, or the interaction, $F(3,32) = 1.24$, $p = .31$, $\eta_p^2 = .10$. Hence, there is no indication that training effects transfer more strongly to operation-inverted problems than to untrained problems.

Although the small standard error bars in Fig. 1 provide a direct visual indication that the sample means in the operation-inverted and untrained problems conditions are likely to be very similar to those of the population, supplementary post hoc power analysis was performed on the null effect of transfer condition in this ANOVA using G*Power 3.1.3 (Erdfelder, Faul, & Buchner, 1996).[2] The power to detect a small effect size of at least $f = .10$ (Cohen, 1988) was .87 and the power to detect a medium effect ($f = .25$) was greater than .99. These high power results may seem surprising. They can be understood, however, in the context of the strong correlation (pooled over age groups) between the number

[1] Analyses were conducted using SAS PROC GLM and Type III sums of squares. Type I and Type III sums of squares yielded nearly identical outcomes, confirming that the slightly unbalanced design with respect to the age factor had negligible impact on the results.

[2] The selected G*Power options for these analysis were as follows: F test, repeated measures ANOVA (within), $\alpha = .05$, $N = 36$, number of age groups = 4, number of repeated measures = 2, correlation = .933.

correct in the inverted problems condition and the number correct in the untrained problems condition ($r = .933$).

Accuracy

Accuracy results are shown in Fig. 1B. ANOVAs identical to those described above were performed, with the same pattern of results. In the ANOVA comparing trained problems with the mean of the other two transfer conditions, there was a highly significant effect of transfer condition, $F(1,32) = 13.04$, $p < .001$, $\eta_p^2 = .29$, but no significant effect of either age or the transfer condition by age interaction, $Fs(3,32) < 1.0$. In the ANOVA comparing the operation-inverted and untrained problems conditions, there were again no significant effect of age, transfer condition, or their interaction.

Pre–Post results

Although the posttest results described above are optimal (in that they include all 36 students of the sample) for testing whether there were *relative* differences in transfer of learning among the conditions, those results do not address the second major question outlined in the Introduction—whether there was any *absolute* transfer of learning to untrained (or operation-inverted) problems. That question can be addressed by evaluating whether there was any pre–post performance improvement for untrained or operation-inverted problems among the subset of 22 students who attended the pretest session.

If there is pre–post improvement for untrained and/or operation-inverted problems, that finding could reflect (a) transfer of learning from the experimentally trained problems and/or (b) learning from school or homework over the 4-week training period. If there is no observed pre–post improvement, we can infer that (a) training in the current experiment transferred minimally, if at all, to inverted or untrained problems and (b) there was minimal or no fluency learning in school for problems from the assigned master sets.

Preliminary analyses indicated no statistically significant differences in the pre–post results as a function of age or operation pair (addition–subtraction vs. multiplication–division), as was expected based on the posttest results described above. Thus, analyses described below were collapsed over those factors.

Number correct

Number correct is plotted in Fig. 2A. Three orthogonal single degree of freedom contrasts were conducted. The first contrast, which tested for a pre–post learning effect limited to the untrained problems, was not significant, $F(1,21) = 2.00$, $p = .17$, $\eta_p^2 = .09$, as was the second contrast, which tested for the possibility of greater pre–post learning in the operation-inverted problems condition than in the untrained problems condition (i.e., the interaction between pre- and posttest and transfer condition, excluding trained problems), $F(1,21) > 1.0$. The third contrast, which tested whether there was more pre–post learning in the trained problems condition than in the other two conditions combined (i.e., the interaction involving trained problems vs. the mean of operation-inverted and untrained problems) was highly significant, $F(1,21) = 30.04$, $p < .0001$, $\eta_p^2 = .59$. All of these results are in accord with the posttest results, and they also demonstrate little or no absolute transfer of learning to the operation-inverted and untrained problems conditions.

Accuracy

Accuracy is plotted in Fig. 2B. All contrasts map to those described above. The first contrast did not approach significance, $F(1,21) < 1.0$, indicating no accuracy gain for untrained problems, nor did the second contrast, $F(1,21) < 1.0$, indicating no difference in accuracy gain between the operation-inverted and untrained problems conditions. The third contrast just reached significance, $F(1,21) = 4.44$, $p < .047$, $\eta_p^2 = .17$, suggesting more pre–post accuracy improvement in the trained condition than in the mean of the operation-inverted and untrained problems conditions. These accuracy results, as well as those of the posttest analyses, eliminate a speed–accuracy trade-off account of the number correct results.

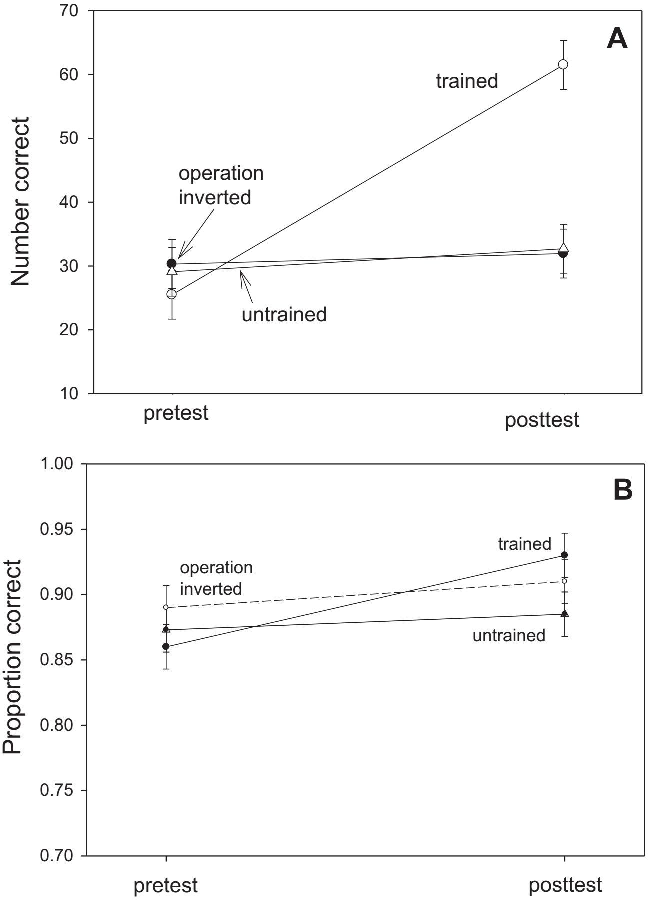

Fig. 2. Pre–Post results for number correct (A) and accuracy (B) plotted by transfer condition. Error bars are standard errors based on the error term of within-participants factorial ANOVAs performed on test (pre or post) and transfer condition (trained, inverted, or untrained).

Discussion

Across all ages and for both pairs of arithmetic operations, the current delayed transfer results agree fully with those for adults; the substantial learning that was observed for trained problems did not transfer to either operation-inverted or untrained problems. Those results appear to exclude operation-level procedure learning as the primary driver of learning among children, at least in the current paradigm. That finding is surprising in light of the aforementioned literature documenting procedural strategy shifts among children. One possible account is that training on a relatively small problem set accelerates the shift to IE-based retrieval and item-level procedural gains relative to operation-level procedural gains. The results also raise the possibility that operation-level procedural shifts are more a consequence of direct instruction (of which there was none in the current experiment) than of independent strategy discovery.

The variant of complement problem mediation described in the Introduction, which predicts cross-operation transfer, is also inconsistent with our results. Rather, the results suggest that mediation either (a) was not used frequently by children in the current experiment or (b) was used and may have contributed to learning but did not yield cross-operation transfer. A model that presumes the latter case would need to posit some mechanism by which the substantial learning that occurred during training was not accessible to support mediation. One possibility is that the learning for addition and multiplication during training was too specific (e.g., tied to format of presentation) to be accessible during subtraction or division mediation on the transfer test. Such an account, however, would need to accommodate evidence that, among adults at least, the majority of improvement with training is not specific to presentation format (Rickard & Bourne, 1996, Experiment 2). In any case, an important goal for future work is development of a mediation model that can integrate findings from transfer, verbal report, and problem format experiments. That goal would be facilitated by development of a computational model of the mechanism of complement problem access and of the influence of presentation format, learning history, and other potentially relevant factors.

Other learning accounts that are consistent with the current results include item-level procedural learning and a shift to (or strengthening of) IE-based memory retrieval. For multiplication–division, performance rate analyses of the posttest data favor IE-based learning over item-level procedural learning for at least a subset of participants. Those rates were estimated for each student by dividing the duration of each posttest block (180 s) by the number of attempted problems. For multiplication–division, the median performance rate over students in the training condition was 2.1 s per problem, which is well within the range of memory-based processing for children (e.g., Lemaire & Siegler, 1995). In contrast, for untrained problems, the median performance rate was 5.6 s, which suggests the use of procedures or of a mixture of procedural and retrieval trials for those problems. Thus, it is reasonable to infer that, for at least half of the students in the multiplication–division group (i.e., those performing at better than the median rate noted above), IE-based learning was a major component of the training gains.

For addition–subtraction, the median performance rate for trained problems was 5.8 s and the median performance rate for untrained problems was 10.0 s. These results do not discriminate cleanly between IE-based learning and item-level procedure learning. However, the five students who performed best on trained problems had a median performance rate of 3.2 s, whereas their median performance rate on untrained problems was 9.0 s. Those results suggest that at least a portion of addition–subtraction learning for some students is likely to have taken the form of IE-based retrieval gains.

Specificity of arithmetic learning across development

For the case of independent (i.e., non-tutored) training explored here, the results do not support the hypothesis that children's arithmetic learning is more flexible (less specific) than that for adults. Rather, the specificity of learning observed in this study is nearly identical to that observed among adults, suggesting that high specificity is a developmental constant, at least in the current training context and for learning that is retained over at least 1 day. A similar high degree of learning specificity among children was observed by Walker, Mickes, Bajic, Nailon, and Rickard (2013). They compared transfer of learning following multiple sessions of drill versus arithmetic fact triangle practice and found that, whereas fact triangle practice promoted better fact triangle performance, that learning did not transfer at all to subsequent arithmetic answer production. It remains to be determined whether that high degree of specificity holds for children in other domains.

Applied implications

From an educational perspective, the current results raise the possibility that the intuitive pedagogical approach of encouraging students to use operation-inverted complement problems to facilitate performance on a newly introduced operation (e.g., encouraging students to think about the multiplication problem to help solve the complement division problem) should be used with caution. Our results suggest that students may fail in that endeavor, delaying the execution of the procedures that eventually generate the answer and, perhaps as a consequence of repeatedly failing to accomplish something that the instructor believes they should be able to do, negatively affecting their arithmetic self-efficacy. The results also raise the possibility that a strategy of emphasizing addition and multiplication training at the expense of subtraction and division would leave children unprepared to advance to the next stage of mathematics. Rather, the observed specificity implies that, given an educational goal of equivalent fluency for all four operations, at least as much training time should be allocated to subtraction and division. Indeed, given the prior empirical support for IE-based memory, along with the fact that the IE model implies that twice as many facts need to be memorized for subtraction and division than for addition and multiplication (Rickard, 2005), substantially *more training time* may be needed to master subtraction and division than to master addition and multiplication. Follow-up research on the specificity of children's arithmetic learning is needed to further evaluate these possibilities.

Finally, the current results apply most directly to the case of independent (i.e., non-tutored) training effects, as may occur in the context of homework, testing, and many classroom activities. Different transfer results may be observed using other training paradigms, particularly those that focus on teaching concepts, reasoning strategies, and special rules and shortcuts. Research is needed that compares both learning rate and transfer in non-tutored versus tutored paradigms.

Acknowledgments

We thank C. Renell Nailon and the Spring Valley Community Center for their cooperation and enthusiasm in conducting this study. We also thank research assistants whose efforts were central to the execution of this study: Jessica Campos, Travis Carlisle, Micha Fernandez, Monica Guzman, Vivian Hwe, Emon Lagevardi, Perla Padilla, Julia Trigeiro, and Yen Vu.

Appendix

Master sets and subsets of number triplets from which stimulus sets were constructed.

Easy addition–subtraction set			Difficult addition–subtraction set		
Subset A	Subset B	Subset C	Subset A	Subset B	Subset C
2 4 6	2 7 9	2 5 7	3 5 8	4 5 9	3 7 10
2 9 11	6 6 12	3 3 6	2 9 11	2 6 8	3 4 7
4 4 8	4 9 13	4 7 11	4 6 10	5 6 11	2 8 10
3 9 12	2 3 5	4 8 12	5 7 12	3 6 9	3 8 11
3 5 8	4 5 9	3 7 10	6 7 13	7 8 15	5 8 13
2 6 8	3 6 9	3 4 7	8 9 17	3 9 12	5 9 14
4 6 10	5 6 11	2 8 10	4 9 13	6 8 14	4 7 11
5 7 12	3 8 11	5 8 13	6 7 16	4 8 12	6 9 15
Easy multiplication–division set			Difficult multiplication–division set		
2 4 8	2 7 14	2 5 10	3 5 15	2 6 12	3 4 12
2 9 18	4 4 16	3 3 9	2 9 18	3 6 18	2 8 16
6 6 36	4 9 36	4 7 28	4 6 24	4 5 20	3 7 21
3 9 27	2 3 6	4 8 32	5 7 35	3 9 27	3 8 24
5 3 15	4 5 20	3 7 21	6 7 42	5 6 30	4 7 28
2 6 12	3 6 18	3 4 12	6 8 48	4 8 32	5 8 40
4 6 24	5 6 30	2 8 16	7 9 63	4 9 36	5 9 45
5 7 35	3 8 24	5 8 40	8 9 72	7 8 56	6 9 54

References

Bajic, D., Kwak, J., & Rickard, T. C. (2011). Specificity of learning through memory retrieval practice: The case of addition and subtraction. *Psychonomic Bulletin & Review, 18*, 1148–1155.
Baroody, A. J. (2010). Children's relational knowledge of addition and subtraction. *Cognition and Instruction, 17*, 137–175.
Barrouillet, P., Mignon, M., & Thevenot, C. (2008). Strategies in subtraction problem solving in children. *Journal of Experimental Child Psychology, 99*, 233–251.
Burns, M. K. (2005). Using incremental rehearsal to increase fluency of single-digit multiplication facts with children identified as learning disabled in mathematics computation. *Education and Treatment of Children, 28*, 237–249.
Campbell, J. I. D. (1999). Division by multiplication. *Memory & Cognition, 27*, 791–802.
Campbell, J. I. D. (2008). Subtraction by addition. *Memory & Cognition, 36*, 1094–1102.
Campbell, J. I. D., & Agnew, H. (2009). Retrieval savings with non-identical elements: The case of simple addition and subtraction. *Psychonomic Bulletin & Review, 16*, 938–944.
Campbell, J. I. D., & Alberts, N. M. (2010). Inverse reference in adults' elementary arithmetic. *Canadian Journal of Experimental Psychology, 64*, 77–85.
Campbell, J. I. D., Fuchs-Lacelle, S., & Phenix, T. L. (2006). Identical elements model of arithmetic memory: Extension to addition and subtraction. *Memory & Cognition, 34*, 633–647.
Campbell, J. I. D., & Robert, N. D. (2008). Bidirectional associations in multiplication memory: Conditions of negative and positive transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 546–555.
Cohen, J. (1988). *Statistical power for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28*, 1–11.
Fendrich, D. W., Healy, A. F., & Bourne, L. E. Jr., (1993). Mental arithmetic: Training and retention of multiplication skill. In C. Izawa (Ed.), *Applied cognitive psychology: Applications of cognitive theories and concepts* (pp. 111–133). Hillsdale, NJ: Lawrence Erlbaum.
Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review, 79*, 329–343.
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.
Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General, 124*, 83–97.
Mauro, D. G., LeFevre, J., & Morris, J. (2003). Effects of problem format on division and multiplication performance: Division facts are mediated via multiplication-based representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 163–170.
Peters, G., De Smedt, B., Torbeyns, J., Ghesquire, P., & Verschaffel, L. (2010). Adults' use of subtraction by addition. *Acta Psychologica, 135*, 323–329.
Rickard, T. C. (2005). A revised identical elements model of arithmetic fact representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 250–257.
Rickard, T., & 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.
Rickard, T. C., Healy, A. F., & Bourne, L. E. Jr., (1994). On the cognitive structure of basic arithmetic skills: Operation, order, and symbol transfer effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1139–1153.
Robinson, K. M. (2001). The validity of verbal reports in children's subtraction. *Journal of Educational Psychology, 93*, 211–222.
Robinson, K. M., Arbuthnott, K. D., Rose, D., McCarron, M. C., Globa, C. A., & Phonexay, S. D. (2006). Stability and change in children's division strategies. *Journal of Experimental Child Psychology, 93*, 224–238.
Rusconi, E., Galfano, G., Rebonato, E., & Umiltà, C. (2006). Bidirectional links in the network of multiplication facts. *Psychological Research, 70*, 32–42.
Shrager, J., & Siegler, R. S. (1998). A model of children's strategy choices and strategy discoveries. *Psychological Science, 9*, 405–410.
Siegler, R. S., & Jenkins, E. A. (1989). *How children discover new strategies*. Mahwah, NJ: Lawrence Erlbaum.
Walker, D., Mickes, L., Bajic, D., Nailon, C., & Rickard, T. C. (2013). A test of two methods of arithmetic fluency training and implications for educational practice. *Journal of Applied Research in Memory and Cognition, 2*, 25–32.