Does Fostering Reasoning Strategies for Relatively Difficult Basic Combinations Promote Transfer by K–3 Students?

Article *in* Journal of Educational Psychology · January 2015

DOI: 10.1037/edu0000067

5 authors, including:

Arthur J. Baroody
University of Illinois, Urbana-Champaign

David J Purpura
Purdue University

Michael Eiland
University of Illinois, Urbana-Champaign

Erin E Reid
Erikson Institute

Journal of Educational Psychology
2016, Vol. 108, No. 4, 576–591

0022-0663/16/$12.00 http://dx.doi.org/10.1037/edu0000067

Does Fostering Reasoning Strategies for Relatively Difficult Basic Combinations Promote Transfer by K-3 Students?

Arthur J. Baroody
University of Illinois at Urbana-Champaign

David J. Purpura
Purdue University

Michael D. Eiland
University of Wisconsin

Erin E. Reid
Erikson Institute, Chicago, Illinois

Veena Paliwal
University of West Georgia

How best to promote fluency with basic sums and differences is still not entirely clear. Some advocate a direct approach—using drill to foster memorization of basic facts by rote. Others recommend an indirect approach that first involves learning reasoning strategies. The purpose of the present study was to evaluate the efficacy of 2 computer-based interventions that embody an indirect approach by highlighting the conceptual bases for 2 relatively difficult reasoning strategies: subtraction as addition (e.g., for $8 - 5$, think: "What plus 5 equals 8?") and use-10 (e.g., "If $10 + 5 = 15$ and 9 is 1 less than 10, then $9 + 5$ is 1 less than 15"). After pretest, 85 Grade K-3 students were randomly assigned to subtraction, use-10, or drill conditions. The subtraction and use-10 conditions served as an active control and represented regular classroom instruction for each other; the drill condition controlled for the effect of extra practice and represented the direct approach. Each intervention involved 30-min sessions, twice weekly, for 12 weeks. Using pretest scores, mathematics achievement, and age as covariates, mixed-model analyses of covariance revealed that, at posttest at least 2 weeks later, the subtraction group outperformed both comparison groups on progress toward fluency and fluency rate with unpracticed subtraction items. The use-10 group achieved analogous results with unpracticed add-with-8 or -9 combinations. This transfer suggests that the conceptually based indirect programs were efficacious and more successful than regular classroom instruction or the direct approach in promoting progress toward fluency and fluently itself.

Keywords: elementary math, basic addition and subtraction facts, instructional design, learning theory, conceptual change

Supplemental materials: http://dx.doi.org/10.1037/edu0000067.supp

In a guest editorial of a local paper, a parent, who was upset by the changes in the elementary mathematics curriculum prompted by the Common Core State Standards, wrote, "A father at the PTO meeting . . . said, 'How does my first grader learn math without the fundamentals of learning simple math facts?' He went on to state, 'If you asked any one of us in the room what 2 plus 2 equals, we would say 4. Why? We were taught in the 1st and 2nd grade, the facts, by rote memorization'" (Elliott, 2014, p. 9).

Clearly, primary-grade students need to generate the sums of basic addition combinations, such as $9 + 7 = 16$, and the differences of related subtraction items, such as $16 - 7 = 9$, *fluently*—that is, accurately, quickly, appropriately, and adaptively (Council of Chief State School Officers [CCSSO], 2010; National Council of Teachers of Mathematics [NCTM], 2000, 2006; National Mathematics Advisory Panel [NMAP], 2008; National Research Council [NRC], 2001). However, many students struggle to achieve this goal (Henry & Brown, 2008), particularly with larger sums and differences (see reviews by Baroody, Bajwa, & Eiland, 2009; Cowan, 2003; NMAP, 2008). Unfortunately, it remains unclear how instruction can best promote fluency with more difficult basic combinations.

This article was published Online First October 12, 2015.

Arthur J. Baroody, College of Education, University of Illinois at Urbana-Champaign; David J. Purpura, Human Development and Family Studies, Purdue University; Michael D. Eiland, Wisconsin Center for Education Research, University of Wisconsin-Madison; Erin E. Reid, Early Math Collaborative, Erikson Institute, Chicago, IL; Veena Paliwal, Department of Mathematics, University of West Georgia.

The research reported herein was supported by the Institute of Education Science, U.S. Department of Education, through Grant R305A080479 ("Fostering Fluency with Basic Addition and Subtraction Facts"). Preparation of this article also was supported, in part, by the U.S. Department of Education (Grant R305B100017: "UIUC Postdoctoral Research Training Program in Mathematics Education"). The opinions expressed are solely those of the authors and do not necessarily reflect the position, policy, or endorsement of the Institute of Education Science or the U.S. Department of Education.

Correspondence concerning this article should be addressed to Arthur J. Baroody, College of Education, University of Illinois at Urbana-Champaign, Champaign, IL 61820. E-mail: baroody@illinois.edu

Elliott's (2014) editorial reflects the long-standing conventional wisdom among many policymakers, educators, parents, and cognitive psychologists that the 121 basic sums and the 66 basic differences can be taught most effectively by using massive amounts of drill to promote directly the memorization of these facts by rote. In contrast to this "direct-facts approach," the Common Core State Standards reflect a reform-based view that the memorization of basic combinations can be better achieved by a more indirect approach—by first teaching general and meaningful strategies. For example, Common Core Grade 1 goals include the following: "Add and subtract within 20 [using] strategies such as . . . making ten (e.g., $8 + 6 = 8 + 2 + 4 = 10 + 4 = 14$)" and viewing "subtraction as unknown-addend problem [such as] subtract $10 - 8$ by finding the number that makes 10 when added to 8" (hereafter called the *subtraction-as-addition strategy*). Such reasoning strategies enable students to use known knowledge of an easy sum and logical thinking to deduce an unknown and difficult sum or difference, instead of simply guessing or using a laborious and, perhaps error-prone, counting strategy. Proponents of an indirect approach contend that learning and applying reasoning strategies more effectively fosters fluency with relatively difficult basic combinations than a direct-facts approach (Rathmell, 1978).

The purpose of the present study is to evaluate the efficacy of two computer-assisted instructional programs that embody an indirect approach with Grade K through 3 students. The use-10 program focused on using a relatively easy add-with-10 combination to determine logically the sum of a relatively difficult add-with-8 or -9 item (e.g., if $10 + 5 = 15$ and 9 is 1 less than 10, then $9 + 5$ is 1 less than 15); the subtraction program focused on using a relatively easy sum to deduce its relatively difficult, related difference. The efficacy of each indirect program was evaluated by comparing it to an active control that represented regular classroom instruction and—given the competing views about how best to foster fluency with basic combinations—a drill program that embodied the direct-facts approach. The drill program provided repeated, unstructured practice with basic add-with-8 or -9 and subtraction items. Discussed, in turn, are the rationale for instructional approaches that focus on learning factual knowledge, an indirect approach that focuses on relational knowledge, and the present study.

Approaches That Focus on the Goal of Promoting Factual Knowledge

Proponents of the direct-facts approach have argued that an indirect approach is unnecessary for memorizing basic arithmetic facts, a barrier to this process, too difficult for students, or a less efficient use of limited instructional time than using drill (Bezuk & Cegelka, 1995; Goldman & Pellegrino, 1986; Hasselbring, Goin, & Bransford, 1988; Koscinski & Gast, 1993a, 1993b; Silbert, Carnine, & Stein, 1990). Indirect methods have long been viewed as a crutch on which students may become overly dependent and, thus, a disruption to the real business of memorizing the basic sums and differences by rote (see Baroody, 1985, and Cowan, 2003, for reviews). Summarizing her own and others' research, Steinberg (1985) concluded that, even for second graders, using known facts to deduce unknown basic combinations was difficult, and using the relations between addition and subtraction to reason out differences was particularly difficult to understand. Walker, Mickes, Bajic, Nailon, and Rickard (2013) found that practice translating fact triangles into equations (see the Appendix), a method commonly used by elementary curricula to promote recognition of the complementary relation between addition and subtraction and fluency with sums and differences, was significantly less effective in improving efficiency with practiced subtraction items than unstructured drill and recommended that the former be de-emphasized in favor of the latter.

In contrast to the champions of the direct-facts approach, proponents of an "indirect-facts approach" take the more moderate view that indirect instruction can be a useful means in achieving the end of memorizing the basic facts. Specifically, most cognitive psychologists now allow that initially encouraging students to use counting and reasoning strategies provides a meaningful basis for practicing sums and differences and reinforcing correct associations. However, although reasoning strategies are a more efficient backup than counting strategies when fact recall fails, reasoning strategies are viewed as *inferior* to—more error prone and slower than—efficient fact recall (NMAP, 2008). Thus, the goal of instruction for both advocates of the direct-facts and indirect-facts approaches is replacing inferior indirect strategies—including reasoning strategies—with efficient fact recall (Ashcraft & Guillaume, 2009; NMAP, 2008; Siegler, 1996).

Models of Fact Memorization

Direct-facts and indirect-facts approaches are based on the assumption that memorizing a basic fact entails the relatively simple process of forming and strengthening an association between an expression and its answer and that practice is the basis for increasing associative strength (Ashcraft, 1992), or at least, the most important factor in this process (Siegler, 1987; Torgesen & Young, 1983). Large doses of practice are justified by Thorndike's (1922) law of frequency: The more two stimuli are presented together, the stronger their association (see also Ashcraft, 1992; Campbell, 1987; Logan, 1991; Siegler & Shipley, 1995). For example, according to the distribution-of-associations model (Siegler & Jenkins, 1989) and its successors (Shrager & Siegler, 1998; Siegler & Araya, 2005), a memory trace is laid down each time an item is practiced, and thousands of such traces are necessary to achieve efficient fact recall.

The Role of Transfer

Transfer plays little or no role in popular cognitive models of basic fact memorization and recall. Specifically, the associative recall system for factual knowledge is hypothesized to exclude conceptual and procedural knowledge, such as relational knowledge, general rules, or reasoning strategies, and the recollection of basic facts is achieved independently of nonfactual knowledge (Campbell & Beech, 2014; Campbell & Graham, 1985; Campbell & Therriault, 2013). According to the distribution-of-associations model, for example, the metacognitive and associative systems are "representationally encapsulated"—"neither system can . . . modify the contents of the other" (Crowley, Shrager, & Siegler, 1997, p. 480). Internalizing an arithmetic regularity does not change or otherwise affect the associative recall system, and each basic fact accumulates associative strength independently of other, even related, combinations. For instance, practice with $9 + 7$ has no effect on the distribution of associations of its commuted partner 7 + 9 (Shrager & Siegler, 1998). Knowledge that addition and subtraction are related influences the distribution of associations of a subtraction combination only initially and, then, only minutely. For example, prior knowledge that 3 + 5 = 8 initially gives 3 a slight advantage over other known numbers for the subtraction problem 8 − 5, but thereafter, only extensive and accurate practice strengthens the association between 8 − 5 and 3 to the point where the correct difference is stated reliably (Siegler, 1987).

Rickard's (2005) identical elements model allows for limited transfer within addition from one item to its commuted partner but not between adding with 10 and adding with 8 or 9 or between addition and subtraction. Strengthening of an association is hypothesized to occur when the stimulus elements of a practiced combination, including the operation, match exactly a mental representation. For example, practicing 3 + 5 = 8 would match the mental representation of 3, 5, + → 8 and thus would strengthen the association between 8 and 3 + 5 or its commuted counterpart 5 + 3. However, such practice would not match the representation of 8, 5, − → 3 or 8, 3, − → 5 and thus not strengthen the associations for 8 − 5 = 3 or 8 − 3 = 5.

Supporting evidence. Consistent with the identical elements model, Walker et al. (2013) and Walker, Bajic, Mickes, Kwak, and Rickard (2014) found that unstructured practice improved elementary-level students' long-term efficiency with practiced basic sums but not with unpracticed subtraction complements. For example, a gain in the efficiency of 4 + 7 did not impact the efficiency of 11 − 4. Similarly, Campbell and Beech (2014) found that practice with 0 + *n* generalized to unpracticed combinations involving 0, but that practice with non-0 items did not facilitate college students' efficiency with unpracticed non-0 items and concluded that, with the exception of the add-0 rule, adults normally use a fact recall system for basic sums.

However, the evidence supporting the direct-facts or indirect-facts approach is not clear-cut. Walker et al. (2013, 2014) or Campbell and Beech's (2014) results, for example, are inconsistent with findings indicating that (a) practice with sums can facilitate the efficient recall of related but unpracticed subtraction combinations (Baroody, Purpura, Eiland, & Reid, 2014; Barrouillet, Mignon, & Thevenot, 2008; Campbell & Agnew, 2009; Campbell, Fuchs-Lacelle, & Phenix, 2006); (b) interventions to promote the number-after rule for adding 1 (the sum of 1 and 9 is the number after "nine" in the counting sequence) can significantly improve primary-grade students' fluency with unpracticed add-1 items (Baroody, Eiland, Purpura, & Reid, 2012, 2013; Baroody, Purpura, Eiland, & Reid, 2015); and (c) children and adults alike may use automated procedures with the easiest basic sums (Barrouillet & Thevenot, 2013). Indeed, the appreciably faster response times on unpracticed add-1 items than on either the doubles or other small combinations indicate Campbell and Beech's (2014) participants used a different process—perhaps an automated number-after rule for adding 1. Moreover, their methodology may not have been sufficiently reliable or sensitive to differentiate between reasoning or counting strategies and fact recall. For instance, a floor effect may have prevented a significant decrease in the response time of add-1 items, and data were averaged over trials in which different participants, or even the same participant, used different strategies.

Implications for the Present Study

According to proponents of the direct-facts or indirect-facts approach, basic add-with-8 or -9 or subtraction combinations are relatively difficult to recall because they are practiced relatively infrequently (Ashcraft, 1992; Thorndike, 1922). In the absence of a mechanism for transfer, the recommended solution for fostering the memorization of these relatively difficult basic combinations is substantial practice of each fact (NMAP, 2008; Walker et al., 2013, 2014). Equal amounts of modest practice, whether provided by unstructured drill or rehearsing a reasoning strategy should have the same modest impact on the recall of practiced items and no impact on that of unpracticed combinations. Although indirect-facts models allow for transfer due to a reasoning strategy, it is unclear whether modest practice with a reasoning strategy would promote fluent transfer. If it takes thousands of repetitions to achieve fluency with a basic arithmetic fact, presumably it would take more than modest practice to achieve fluency with a relatively complex reasoning procedure.

An Indirect Approach That Focuses on the Goal of Relational Knowledge and Reasoning

Although memorizing basic sums and differences by rote or without the benefit of transfer is possible, it makes learning basic sums and differences more difficult than is necessary (Baroody et al., 2009). Before finally achieving efficiency with basic combinations, students are often highly forgetful and prone to recall errors such as associative confusions, which further reinforces incorrect associations and increases the amount of practice needed to learn a correct association. Additionally, the extensive drill and timed tests often recommended to achieve the memorization of basic facts can be tedious, unpleasant, or overwhelming for many students. Such experiences can create affective barriers to achieving fluency, such as disinterest in mathematics or even math anxiety (Boaler, 2009, 2015; Carpenter, Fennema, Peterson, Chiang, & Loef, 1989; Ramirez, Gunderson, Levine, & Beilock, 2013; Tobias, 1993).

Whereas the direct-facts and indirect-facts approaches are based on the developmental view that the novice-expert progression involves replacing relatively inefficient strategies of counting and reasoning with relatively efficient factual recall that incorporate little or no relational knowledge (Ashcraft & Guillaume, 2009; NMAP, 2008; Siegler, 1996), an "indirect-relations approach" is based on the developmental view that relational knowledge plays a key role in the process of constructing the retrieval system for basic combinations by students and representing and retrieving basic sums and differences by experts (Baroody & Tiilikainen, 2003; Barrouillet et al., 2008; Barrouillet & Thevenot, 2013). Learning relations among facts and using such knowledge to construct reasoning strategies fosters number sense. Such number sense underlies the meaningful memorization of basic combinations, efficiently retrieving this knowledge, and transfer of fluency (Baroody, 1985; Baroody et al., 2009; Brownell, 1935; Jordan, 2007). It also promotes general problem-solving ability, general reasoning ability, and transfer to new content learning. For all of these reasons, nurturing relational learning and reasoning strategies are primary instructional goals in themselves.

Models of Meaningful Memorization

The meaningful memorization of a basic combination or a family of combinations typically involves three overlapping phases (Baroody, 1985; NRC, 2001; Steinberg, 1985; Verschaffel, Greer, & De Corte, 2007). In Phase 1, students use counting strategies to determine an answer. In Phase 2, they use the patterns and relations discovered in Phase 1 to invent reasoning strategies, which they apply in a deliberate manner—consciously and relatively slowly. In Phase 3, students achieve fluent retrieval—that is, they can efficiently, appropriately, and adaptively produce sums and differences from a memory network.

Phase 2 can serve as a bridge between the relatively inefficient counting strategies of Phase 1 and the fluent retrieval of Phase 3 in two critical ways. One way is that an instructional focus on reasoning strategies can highlight relational knowledge, which in turn, can provide an organizing framework for learning, storing, and retrieving both practiced and unpracticed combinations (Canobi, Reeve, & Pattison, 1998; Dowker, 2009; Rathmell, 1978; Sarama & Clements, 2009). A second way is that reasoning strategies about relations can become automatic and incorporated into the retrieval system as a basis of fluent retrieval on a coequal basis with factual recall (Baroody & Varma, 2006; Verschaffel et al., 2007).

The Role of Transfer

Transfer due to relational knowledge and reasoning strategies plays a vital role in models of meaningful memorization. For example, practice with complementary sums and applying the subtraction-as-addition strategy may produce *complement problem mediation*—that is, improved fluency with related, but unpracticed, subtraction combinations (Baroody, 1999; Campbell, 2008; Peters, De Smedt, Torbeyns, Ghesquière, & Verschaffel, 2010).

Supporting Evidence

Although teaching reasoning strategies has long been recommended as a means of transitioning from Phase 1 to Phase 3 (Rathmell, 1978; Suydam & Weaver, 1975; Thornton, 1978), there are surprisingly few methodologically sound evaluations of an indirect approach or comparisons of an indirect approach with the direct-facts approach. A few studies reported instruction that encourages students to learn reasoning strategies is more effective than drill in facilitating retention or transfer of basic combinations (for reviews, see Baroody, 1985; Baroody & Tiilikainen, 2003; Brownell, 1941; Eiland, 2014). For example, Henry and Brown (2008) found that, whereas the use of textbooks that focused on memorizing all basic addition and subtraction facts by rote and timed tests were negatively related to learning basic combinations and the use of flashcards had no positive effect, teaching reasoning strategies was positively correlated with fluency gains at the end of first grade. However, the accumulated research has a number of methodological limitations (see Eiland, 2014). For example, Thornton (1978) found that relating harder combinations to easier ones produced better retention than drill, but she apparently did not examine transfer. Other limitations include not using a reaction time (RT) criterion or otherwise controlling for responses generated by counting or slow reasoning processes or not assessing treatment fidelity (Henry & Brown, 2008; Thornton, 1978).

Efforts to evaluate the efficacy of teaching the subtraction-as-addition strategy or the complement problem mediation hypothesis are especially sparse and have had little or no success. Walker et al. (2013) found that the indirect approach did not promote transfer to unpracticed items. Baroody (1999) found that even explicitly and repeatedly suggesting the subtraction-as-addition strategy over 24 sessions neither helped 6.0- to 7.5-year olds grasp that addition can serve as a shortcut for determining related differences nor produced significantly better transfer to unpracticed differences than did the same amount of drill with practiced sums and differences. Follow-up studies that included providing a rationale for the subtraction-as-addition strategy failed to produce significant transfer to unpracticed subtraction combinations by second graders (Baroody, 2013) and third graders (Baroody, Eiland, & Baroody, 2011b), but the modest effect size in the latter study indicated the intervention was promising at least for older primary-grade students. Similarly, a revised program had a nonsignificant, but promising, impact on first graders' fluency with unpracticed subtraction items (Baroody et al., 2014). In regard to the small amount of other prior evidence supporting complement problem mediation (Barrouillet et al., 2008; Campbell & Agnew, 2009; Campbell et al., 2006), Walker et al. (2014) noted that it is unconvincing because of methodological limitations.

Four reasons may account for the limited positive results of previous intervention efforts. One is that many participants in previous training studies may not have been fluent with related sums. A second reason is that the direct-facts approach of simply practicing basic sums or differences with no explicit guidance, as was done by Walker et al. (2014), or the indirect approach of merely translating fact triangles into equations, as was done by Walker et al. (2013), is not likely to help students understand or invent the nonobvious subtraction-as-addition strategy (Canobi, 2009; Putnam, deBettencourt, & Leinhardt, 1990). Baroody's (1999) indirect intervention—prompting the use of the strategy without helping participants understand its conceptual rationale—may not have been meaningful to participants. Even Baroody et al.'s (2011b; 2014) follow-up studies, which involved an attempt to help participants understand the rationale of the subtraction-as-addition strategy, may have fallen short of the goal of an indirect-relations approach—achieving the relational knowledge needed for a conceptual understanding of the strategy. For example, feedback was provided on correctness but not on why a particular response was correct or incorrect. A third reason is that timed-practice drills as used in Walker et al. (2013, 2014) are not likely to be engaging or motivating to many students. A fourth reason is that, even if an intervention is successful in helping students understand and learn the subtraction-as-addition strategy, it may take distributed practice over many months, rather than weeks as with the Walker et al. (2013) study, to automatize the strategy.

Similarly, the few previous intervention experiments that focused on fostering fluency with add-with-8 or -9 items have had little or no success. Efforts to impose the complex and difficult-to-understand making-10 strategy, even if supported with an effort to make the rationale known, are not likely to result in relational learning and, thus, fluency with the strategy. Murata (2004) found that Japanese students taught the making-10 strategy with larger-addend-first combinations did not apply the strategy when smaller-addend-first combinations were introduced. Four previous intervention efforts by Baroody (2013; Baroody, Eiland, Bajwa, &

Baroody, 2009; Baroody, Eiland, & Baroody, 2011a, 2011b) also failed to produce significant transfer to unpracticed add-with-8 or -9 combinations.

The making-10 strategy may be difficult to learn or apply successfully because one or more of its three components can be challenging for students. Specifically, a student must be fluent with the following components so as not to overtax working memory: (a) decomposing numbers to 9, (b) the associative law of addition and recomposing to make 10 via the number-after rule for adding 1, and (c) the rule that a single-digit n added to 10 results in the sum n + teen. For 9 + 7, for example, this strategy entails (a) decomposing the 7 into 1 and 6; (b)—while holding the 6 in working memory—adding the 1 to 9 to make 10; and (c) finally adding 10 and 6. However, Murata (2004) suggested many students might have difficulty with the first two steps. Furthermore, Canobi et al. (1998) found that first and second graders reported using the relatively arcane associative strategy as a computational shortcut on only 11% of the applicable problems, and students up through Grade 8 do not do much better (Robinson & Dubé, 2009; Robinson, Ninowski, & Gray, 2006).

As many primary-level students struggle to learn or apply the making-10 strategy, Baroody et al. (2014) developed and evaluated an intervention for promoting the use-10 strategy, which does not require decomposition, associativity, recomposition to 10, or retaining the remaining portion of decomposed number in working memory. For 9 + 7, for instance, use-10 entails knowing that 10 + 7 = 17 (Premise 1); knowing that the number before another in the counting sequence is 1 less than its successor (e.g., 9 is 1 fewer than 10; Premise 2); and deducing that the sum of 9 + 7, therefore, must be 1 less than that of 10 + 7 = 17, which is the number before 17, namely 16. The evaluation of the earlier version of the use-10 program (Baroody et al., 2014) produced promising but nonsignificant transfer to unpracticed items.

Implications for the Present Study

According to proponents of the indirect-relations approach, basic combinations involving addition with 8 or 9 or subtraction are relatively difficult to retrieve because instruction inadequately focuses on relational knowledge (Baroody et al., 2009). A key for achieving fluency with relatively difficult sums and differences, then, is helping students construct the relational knowledge that underlies reasoning strategies, which can be become an integral part of the automatic retrieval system and efficiently applied to practiced or unpracticed combinations. Contrary to direct-facts or indirect-facts models, even modest practice may enable students with a high degree of developmental readiness to achieve Phase 3 and exhibit relatively fast and accurate reasoning-based transfer with unpracticed items related to known sums.

Rationale for the Present Study

Although previous attempts to teach subtraction-as-addition and use-10 strategies yielded limited or no positive results, a well-designed indirect-relations approach might be more successful than regular classroom instruction or a direct-facts approach in promoting the learning these general strategies and their fluency.

Features of the Revised and Highly Guided Interventions

The revised subtraction and use-10 programs retained the promising features detailed in Baroody et al. (2014). However, to more effectively implement an indirect-relations approach, the following refinements to these programs were made to better underscore relevant relations and, thus, improve the chances of promoting the learning of general reasoning strategies and transfer:

1. A new Clue game replaced the less interesting and ineffective Wall Help game. The subtraction version of clue entailed explicitly asking, for instance, which of the following five addition items would help answer 9 − 5: 2 + 5, 4 + 5, 4 + 6, 4 + 7, 5 + 6, or 5 + 7. The use-10 version asked, for example, which of the following five add-with-10 combinations would help answer 9 + 6: 5 + 10, 6 + 10, 8 + 10, 10 + 4, or 10 + 6.

2. The Train game, which for the subtraction intervention involved recognizing that the boxcars labeled "4 + 5" and "9 − 5" shared "the same track" and, by way of analogy, the same parts and whole, was significantly improved. For example, the instructions better explained the demands of the task, and a T was added to better underscore the part-whole relations of a fact triangle by separating the whole placed above the T from the parts and the parts, which were placed on each side of the stem of the T, from each other.

3. Feedback more effectively explained why a particular response was correct or not.

Overview of the Present Study

The subtraction group, which received guided intervention on the subtraction-as-addition strategy, and the use-10 group, which received guided intervention on the use-10 strategy, were compared to each other and a third group, which received unstructured practice with both subtraction and add-with-8 or -9 items (drill group). The evaluation of the programs was more methodologically rigorous than that of Baroody et al. (2014) in that participants were randomly assigned within class for all schools and assessed at posttest by testers blind to intervention condition. The present study also included students from Grades K to 3, instead of only those from Grade 1, to extend external validity.

Two measures served to gauge the impact of the indirect-relations interventions. The rate of fluent retrieval used previously by Baroody and colleagues (2012, 2013, 2014), hereafter referred to as *fluency rate*, served to gauge achievement of Phase 3. This measure has ecological validity because the ultimate goal of instruction is fluency. However, an intervention program may be partially successful if participants learn a reasoning strategy and use it accurately, appropriately, and adaptively, if not quickly (i.e., achieve Phase 2, if not Phase 3). In order to provide a more complete picture of the impact of the intervention programs, a new measure, namely a *fluency index*, hereafter referred to as the F-index, was developed to gauge overall progress toward retrieval fluency by taking into account use of Phase 2, as well as Phase 3, strategies. For both dependent measures and both subtraction and add-with-8 or -9 items, retention was assessed with items practiced during the intervention and—more importantly—transfer was assessed with unpracticed items. Transfer is important because it indicates adaptive and appropriate use of a general strategy and is "a primary goal of education" (Kaminski, Sloutsky, & Heckler, 2013, p. 14).

Hypotheses

Hypothesis 1 (H1): Efficacy of the subtraction intervention. The subtraction intervention, redesigned to provide a sounder conceptual basis for the subtraction-as-addition reasoning strategy by more clearly highlighting the relations between subtraction and addition, should facilitate the meaningful learning and memorization of this general strategy above and beyond instruction on subtraction provided by regular classroom instruction, as represented by the use-10 condition. This is because the instruction on the subtraction strategy provided by regular curricula falls far short of a well-designed, indirect-relations approach—is relatively narrow in focus, superficial, and brief—and, thus, likely to be ineffective (Baroody, 2016). The subtraction intervention should also be more efficacious than a direct-facts approach, as represented by the drill condition. This is because even though the drill condition involves supplementing regular classroom instruction with extra, unstructured subtraction practice, such practice is not likely to prompt spontaneous discovery of the nonsalient subtraction-as-addition strategy. As a result, at posttest, the subtraction group should have a significantly better mean F-index score and fluency rate than either comparison group for the practiced and unpracticed subtraction combinations.

Hypothesis 2 (H2): The efficacy of the use-10 intervention. For reasons similar to those given for H1, the use-10 group should exhibit parallel results with practiced and unpracticed add-with-8 or -9 combinations.

Method

Participants

Participants were recruited from 23 K-3 classes in three elementary schools in two school districts serving two medium-size Midwestern U.S. cities. Parental consent forms were returned for 190 students. Ten students were lost to attrition before subject assignments were made due to a school suspension ($n = 1$), a prolonged family vacation ($n = 1$), family relocation ($n = 6$), or autism ($n = 2$). A screening test, which served to identify mastery of the most basic sums and differences, identified a total of 88 students who were at least developmentally ready to learn more advanced reasoning strategies. Mental-arithmetic pretesting identified three students who had already mastered most of the add-with-8 or -9 and subtraction pretest items and were excluded from the study. The 85 students identified as eligible for the study were randomly assigned to an intervention condition. Attrition due to family relocation ($n = 3$) or refusal to participate ($n = 1$) resulted in 29, 27, and 25 students in the subtraction, use-10, and practice-control conditions, respectively, who completed the training and posttesting. Of these 81 participants (5.7 to 10.1 years of age; mean & median = 7.9 years; $SD = 0.97$), 33.3% were male, and 66.7% had at least one factor associated with academic achievement difficulties, as determined by project achievement testing or school records and personnel. The final sample was 46.9% African American, 26.0% Caucasian, 11.1% Hispanic; 3.7% Asian; and 12.3% multiracial, unknown, or other race. Additionally, 74.1% of participants were eligible for free or reduced-price lunch. See the online supplemental materials for the demographic information broken down by condition.

The 14 classes in Schools 1 and 2 used *Everyday Mathematics* (University of Chicago School Mathematics Project [UCSMP], 2005). The nine classes in School 3 used *Math Expressions* (Fuson, 2006). Specific details regarding how these two curricula approached mental arithmetic, in general, and subtraction and add-with-8 or -9 items, in particular, can be found in the online supplemental materials.

Measures

Test of Early Mathematics Ability—Third Edition (TEMA-3). A nationally normed test for children age 3 to 8 years, the Test of Early Mathematics Ability—Third Edition (TEMA-3; Ginsburg & Baroody, 2003), was administered to determine if a participant was at risk because of low achievement. The TEMA-3 achievement score also served as a control variable in the data analyses. The TEMA-3 measures informal and formal concepts and skills in the following domains: numbering, number-comparison, numeral literacy, combination fluency, and calculation. Cronbach's alpha for the form used (Form A) overall is .94. The form's test—retest reliability is .83 and its coefficient alphas for males, females, European, African, Hispanic, and Asian Americans are all .98 and for low mathematics achievers, .99. In terms of criterion-predictive validity, correlations between the TEMA-3 and similar measures (Diagnostic Achievement Battery, KeyMath-R/NU, Woodcock-Johnson III, and Young Children's Achievement Test) range from .54 to .91.

Mental-arithmetic screening test. A screening test served to ensure participants were developmentally ready for the present study and involved relatively easy combinations: eight add-0, 16 add-1, 10 doubles up to $12 + 12$ such as $3 + 3$ and $8 + 8$, six other sums of 10 or less such as $3 + 5$ and $7 + 2$, three $n - 1$, and five subtraction complements of the doubles, such as $10 - 5$ and $12 - 6$. See Baroody et al. (2014) for details regarding the specific items and item order used. The same testing procedure as that described for the mental-arithmetic test in the next subsection was used. Only students fluent on more than 50% of add-with-0 and add-with-1 items and 33% of all items on the screening test were then administered the mental-arithmetic pretest.

Mental-arithmetic test. In order to evaluate the impact of the interventions, the mental arithmetic test included seven practiced subtraction items, seven unpracticed subtraction items, nine practiced add-with-8 or -9 items, and eight unpracticed add-with-8 or -9 items. The test also included a fifth category of items: nine addition complements of the practiced and unpracticed subtraction items. The test items were presented in four sets in a random order, except that two combinations with the same numbers or answer were not presented one after another, commuted combinations or a subtraction item and its addition complements were presented in different sets, and the types of combinations were evenly distributed across the four sets. See the online supplemental materials for the specific items in each category; which groups, if any, practiced an item during their intervention; and the makeup of specific sets.

After a practice session with eight nontest items (3 + 8, 10 − 3, 2 + 6, 10 + 9, 9 − 4, 3 + 5, 13 − 2, and 4 + 7), testing entailed two sessions. Each test session involving a test set of 10 items, a computer reward game, a second test set of 10 items, and a second computer reward game. The test sets were presented in the context of a computer game in which children were encouraged to respond accurately and quickly without counting. Items were presented on the computer screen in the same format used in the interventions, namely as a horizontal arithmetic expression. See the online supplemental materials for additional details on procedure.

Data were gathered on strategy, response time, and response accuracy. Testers identified whether a participant used a counting, reasoning, or undetermined strategy. The criterion for counting was evidence of counting fingers or other objects, verbally citing the counting sequence, or subvocally counting accompanied by successive movement of a finger or the eyes. A reasoning strategy was scored if a child spontaneously exhibited evidence of using deductive reasoning (for 9 + 7, e.g., stating, "10 and 7 is 17, so 1 smaller is 16"), and the type of reasoning was noted in order to identify whether it was consistent with a child's intervention. Response time was determined as detailed in the online supplemental materials. Each response was scored as correct if the sum or difference was accurate and not a false positive due to a response bias. For example, consistently stating the subtrahend would result in a false positive for 14 − 7 and 16 − 8. The response bias criteria are detailed in Baroody et al. (2014). Response biases were scored separately for addition and subtraction items. On the pretest, 11 participants were identified as using a response bias on one or both sets. However, there was only one false positive with subtraction items due to a favorite number response bias and none for add-with-8 or -9 combinations. On the posttest, only two participants in the use-10 condition used a response bias and did so on one set only. A favorite number response bias resulted in one and two false positives on one subtraction item and two add-with-8 or -9 items by one of these students and on two subtraction items and one add-with-8 or -9 item by the other.

A research assistant independently scored a child's response and strategy for each item on the pretest and the posttest videotape for each of the seven participants randomly chosen from each of the three conditions. Interrater agreement for strategy was 95.3% on a total of 841 trials and for responses was 98.9%. For each combination category at pre- or posttest, analyses were done for two dependent measures: F-index and fluency rate. The F-index for each group was the mean of its members' F-index. A child's F-index was the mean score of the relevant trials, which were each scored on a 6-point continuum (see the online supplemental materials for the scoring criteria). A computer program scored the F-index. A staff member checked the scoring for a sample of 68% of the 2,258 pretest cases and 11% of 3,259 posttest cases; agreement occurred 98.2% and 95.8% of the time, respectively. All cases where a discrepancy occurred were checked and found to be due to human error. Fluency rate for each group was the mean of its members' proportion of fluent retrieval. A response was scored as fluent and awarded 1 point if it was correct, fast (<3 seconds), and did not involve counting or a response bias. Otherwise, the response was scored as 0 points. The agreement between the scoring by a computer program and a research assistant, both of which involved scoring response biases, was 99.7% for the pretest and 98.2% for the posttest. The practiced and unpracticed items for individual domains were combined at pretest to use as a covariate in the analyses. The combined subtraction task showed acceptable reliability (α = .71) and the combined use-a-10 task showed good reliability (α = .87; Nunnally & Bernstein, 1994).

Computer-Based Instruction

The instructional programs involved five stages. The first two stages were common to all participants and served to prepare them for the computer-assisted mental-arithmetic testing and the experimental instruction. The last three stages varied by group and served as the experimental conditions or interventions. See Table 4 in Baroody et al., 2014, for a more complete description of the aims and plan for each of the five stages.

Preparatory instruction for the study. Stages I and II ensured that all participants could understand the symbolic addition and subtraction expressions and the terminology of the verbal instructions used in the mental arithmetic testing. It also ensured (a) the necessary experiences to use the computerized interventions, such as how to use the virtual manipulatives and enter answers, and (b) the developmental prerequisites, such as solving meaningful word problems concretely, to benefit from interventions.

Experimental interventions. The two indirect-relations programs involved highly guided instruction. Stage III of these programs served to foster a reasoning strategy, and related items were always presented consecutively. For example, in the subtraction condition, the "helper" addition complement 2 + 5 = 7 immediately preceded the targeted subtraction problem 7 − 5 = 2, and in the use-10 condition, the helper add-with-10 item 10 + 7 immediately preceded the targeted add-with-9 item 9 + 7. Stage IV served to practice using a strategy with generous time limits, and related items were often presented successively. Stage V served to practice using a strategy with a stringent time limit to encourage its efficient use, and related items were never presented consecutively. Stages III, IV, and V of the indirect-relations interventions are depicted in the online supplemental materials.

Feedback in Stages III and IV of the indirect-relations programs involved juxtaposing related helper and target equations or expressions and sometimes color-coding to highlight their analogous elements. For example, in the subtraction version of Castle Wall game, the $\underline{9}$ + *5* = **14** and **14** − *5* = $\underline{9}$ equations were aligned one above the other; in the use-10 version, the equation 10 + *6* = 16 was aligned immediately above 9 + *6* = 15. Note that the underlined terms had a common color to indicate an analogous part; the italicized terms had a different color in common to represent another corresponding part; and the bold represents a third color to underscore an equivalent whole. In the Train game for subtraction, the $\underline{9}$ + *5* and **14** − *5* box cars were shown together on the same horizontal track labeled with the fact triangle: **14** / $\underline{9}$ | *5* in order to highlight common their common parts and whole.

All three conditions involved discovery learning in the sense that a participant was not directly told a strategy or what to do. For example, participants in the subtraction intervention were never directly told that to answer 8 − 5 = □, they should determine what must be added to 5 to make 8 or chose the number from $3 + 5 = 8$ that does not appear in $8 - 5 = \square$. All three conditions were game-based and indeed, on the surface, involved the same games. The use-10 condition was analogous to the subtraction intervention, except that it involved relating add-with-8 or -9 items to the related and easier add-with-10 sums. Other commonalities among the three conditions included: (a) the same time limit for responding in Stages III (none), IV (6 seconds), and V (4 seconds); (b) practice with 20 to 24 different items per session; and (c) a standardized item order within a condition. The drill group also practiced the same subtraction items practiced by the subtraction group and the same add-with-8 or -9 items practiced by the use-10 group but in an unordered manner. Moreover, no related helper items were practiced to avoid facilitating relational learning (e.g., $9 + 5 = 14$ was followed by the unrelated subtraction item $7 - 5 = 2$). See the online supplemental materials for a brief description of the instructional games used, which were common to all conditions, and how the games differed by condition.

Fidelity of the Experimental Interventions

Fidelity of all intervention stages was ensured first and foremost by the computer sign-in system, which guaranteed each participant received his or her assigned intervention and completed each lesson without duplication, and the computer programs, which faithfully delivered an intervention in the manner designed. Although a trainer's primary instructional role was to give voice to the scripted instructions and the feedback graphically displayed by the computer screen, staff training, staff oversight, and lesson log sheets further ensured fidelity. In regard to implementation fidelity, all participants completed 100% of the Stage I to V lessons before posttesting. Additional details regarding fidelity are described in Baroody et al. (2014).

Research Design and Procedures

A randomized control trial (RCT) served to evaluate the efficacy of the three programs. Project personnel conducted all project testing and supervised all instruction at a computer station in a hallway outside a child's classroom or in a room dedicated to the project. Pull outs occurred in nonliteracy time blocks, including mathematics instruction and play time. Positive assent was obtained for each testing and instruction session.

During the first 7.5 weeks of the study, all students in the sample pool received the Stage I and II preparatory instruction and simultaneously were tested on the TEMA-3. After the completion of the preparatory instruction, the participants were individually administered the computer-based preliminary mental-addition screening test and then the mental-arithmetic pretest (see Baroody et al., 2014, for details). All participants who met the selection criteria were randomly assigned within class to one of the three interventions. Four classes with a number of participants divisible by three had a balanced assignment. Balancing was maximized in the remaining classes (e.g., one class with five participants had two, two, and one students in the subtraction, use-10, and drill conditions, respectively). The three computer-assisted intervention programs each lasted 12 weeks and were conducted simultaneously. Both the preparatory instruction and the interventions involved two 30-min sessions per week. Instruction and feedback were presented in written form on the screen and verbally by the computer or read by project personnel. The mental arithmetic posttest served to gauge retention and transfer. Retesting of all participants began at least two weeks after the intervention by testers blind to a participant's intervention assignment. Specifically, the first two test sets and the second two test sets were administered 14 to 19 and 15 to 20 days, respectively, after the intervention. The median time, involving 44% of all participants, was 14 days for the first two test sets and 15 days for the last two test sets.

Similarities and differences among the three experimental conditions are summarized in Table 1. The subtraction and use-10 groups practiced mutually exclusive items and, thus, served as an active control for each other. An active-control group controls for various threats to internal validity such as a history effect due, for instance, to regular classroom instruction and practice, a maturation effect, and regression to the mean. Unlike a business-as-usual control, an active control effectively controls for the impact of extra arithmetic instruction and practice, the impact of a novelty or "special treatment" effect, the negative impact on a no-treatment control of not receiving "special treatment," a positive testing effect due to an experimental child's familiarity with research staff, and a negative testing effect due a pure control's relative unfamiliarity with testers. The drill group received unstructured or unordered drill with the subtraction and add-with-8 or -9 items practiced by the subtraction and use-10 groups, respectively, with

Table 1
Similarities and Differences Among the Subtraction, Use-10, and Drill Conditions

Attribute	Condition		
	Subtraction	Use-10	Drill
Regular classroom subtraction instruction and practice	Yes	Yes	Yes
Regular classroom add-with-8 or -9 instruction and practice	Yes	Yes	Yes
Preliminary (Stage I and II) intervention	Yes	Yes	Yes
Experimental (Stage III and IV) intervention on the subtraction-as-addition strategy	Yes	No	No
Supplemental practice (Stages III to V) with practiced subtraction items	Yes	No	Yes
Supplemental practice with unpracticed (transfer) subtraction combinations	No	No	No
Experimental (Stage III and IV) intervention on the use-10 strategy	No	Yes	No
Supplemental practice (Stages III to V) with practiced add-with-8 or -9 combinations	No	Yes	Yes
Supplemental practice with unpracticed (transfer) add-with-8 or -9 combinations	No	No	No

a few exceptions. For instance, in the Does It Help? game, commuted items such as 8 + 7 and 7 + 8 or obviously nonequivalent items such as 7 + 9 and 9 + 9 were presented together. The drill condition served to evaluate whether supplemental practice alone, as prescribed by proponents of the direct-facts or indirect-facts approach, was sufficient to promote fluency with practiced subtraction or add-with-8 or -9 items. As all interventions involved the same computer-based games, differences among the groups cannot be attributed to differences in these experiences.

Analytic Procedure

A 3 (condition: subtraction, use-10, drill) × 2 (experimental experience with items: practiced or unpracticed) mixed-model ANCOVA was conducted separately for the subtraction items and the add-with-8 or -9 items. In each case, condition was a between-groups variable and experimental experience with items (i.e., whether or not a student practiced the items during an intervention) was a within-groups variable. Pretest scores, TEMA-3 achievement score, and age were included as the covariates. Specific contrasts were also conducted to evaluate differences between each group on the practiced and transfer items. Notably, each indirect-relations group was compared to their respective active-control group and the drill group; the drill group and the active-control group were also compared. The Benjamini and Hochberg (1995) adjustment for multiple comparisons was utilized to minimize Type I error. The adjustment was applied separately for subtraction and add-8 or -9 items. For each, there were a total of six comparisons (three comparisons for practiced items and three for transfer items).

Given the importance of reporting effect size (Lipsey et al., 2012; Wilkinson and the APA Task Force on Statistical Inference, 1999), efficacy was also evaluated by using Hedges' *g* for all specific contrasts of interest. Unlike significance levels, effect size bears on practical significance and indicates statistical magnitude of the effect (Lipsey et al., 2012). The Institute of Education Science (IES) supported the research, and per the Institute's What Works Clearinghouse (WWC) guidelines (IES, 2014): "Effects of 0.25 standard deviations or larger are considered to be substantively important . . . even though they may not reach statistical significance in a given study" (p. 23).

Specific contrasts between groups were conducted separately with school, grade, and classroom included as random-effect covariates to check for possible school, grade, and classroom effects. None of these variables were significant factors in any of the analyses. Thus, school, classroom, and grade were not included in the final analyses.

Results

Descriptive statistics for participants' performance for the practiced and unpracticed subtraction and the practiced and unpracticed add-with-8 or -9 combinations are summarized in Table 2 for the F-index and fluency rate. The hypothesized relative performance of the experimental conditions for each type of combination for the F-index and for fluency rate and the actual result as gauged by effect size are reported in Table 3. The preliminary analyses regarding the equivalence of the groups and the analyses of posttest F-index data regarding H1 and H2 are discussed in turn. With

Table 2
Pretest and Adjusted Posttest F-Index and Fluency Rate Mean Scores and Standard Deviations by Condition

	F-Index (0 to 5)				Fluency rate (0 to 1)			
	Pretest		Adjusted[a] Posttest		Pretest		Adjusted[a] posttest	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Practiced subtraction combinations (subtraction items practiced by subtraction and drill conditions)								
Subtraction	.55	.55	2.30	1.09	.06	.12	.32	.22
Use-10	.68	.59	1.11	1.23	.11	.17	.17	.24
Drill	.63	.70	1.86	1.15	.09	.11	.31	.23
Unpracticed subtraction combinations (subtraction transfer items not practiced in any condition)								
Subtraction	.56	.64	1.44	.84	.05	.06	.17	.15
Use-10	.87	.97	.66	.85	.06	.08	.08	.10
Drill	.79	.67	.81	.67	.07	.12	.08	.11
Practiced add-with-8 or -9 combinations (add-with-8 or -9 items practiced by use-10 and drill conditions)								
Subtraction	.57	.60	1.27	.72	.05	.08	.12	.12
Use-10	1.20	1.02	2.10	1.13	.15	.15	.32	.23
Drill	.85	.94	2.22	1.08	.10	.17	.35	.21
Unpracticed add-with-8 or -9 combinations (add-with-8 or -9 transfer items not practiced in any condition)								
Subtraction	.39	.56	.96	.57	.02	.05	.06	.08
Use-10	.77	.80	1.36	1.14	.05	.09	.15	.21
Drill	.54	.79	1.11	.74	.04	.09	.10	.13

[a] Posttest scores adjusted for pretest score. F-index = fluency index.

one exception, all significant comparisons remained significant after application of the Benjamini-Hochberg correction. As the analyses of posttest fluency rate data regarding H1 and H2 largely paralleled those of the F-index, the fluency rate results are detailed in the online supplemental materials.

Preliminary Analyses

Preliminary analyses revealed that the three groups did not differ in gender, $\chi^2(2, N = 81) = 3.01$, $p = .222$; ethnicity (Caucasian, African American, Hispanic, Multiracial, and other), $\chi^2(8, N = 81) = 5.74$, $p = .676$; or age, $F(2, 78) = 0.16$, $p = .853$. An ANOVA did not reveal statistically significant group differences in mathematics achievement, as measured by TEMA-3 standard scores, $F(2, 78) = 0.16$, $p = .851$. ANOVAs revealed no statistically significant differences among groups on three key combination categories at pretest: practiced subtraction, $F(2, 78) = 1.21$, $p = .304$, unpracticed subtraction, $F(2, 78) = 0.35$, $p = .703$, and unpracticed add-8 or -9 items, $F(2, 78) = 1.93$, $p = .151$. Although an ANOVA did reveal statistically significant group differences for the practiced add-8 or -9 items, $F(2, 78) = 3.76$, $p = .028$, including pretest performance as a covariate in the main analysis described for H2 below controls for any initial differences. All main analyses met the assumptions for ANCOVA, including homogeneity of regression.

Table 3
Hypothesized and Actual Comparisons (in Terms of Effect Size) Among Conditions on F-Index and Fluency Rate Scores

Type of combination	Hypothesis	F-Index: Actual results	F-Index: g [95% CI]	Fluency rate: Actual results	Fluency rate: g [95% CI]
Practiced subtraction	SUB ≫ U10	SUB ≫ U10***	***1.00*** [.69, 1.13]	SUB ≫ U10**	***.65*** [.56, .74]
(Subtraction practiced	SUB > DRL	SUB > DRL+	***.39*** [.09, .69]	SUB ≈ DRLns	.04 [.02, .10]
by SUB & DRL conditions)	DRL > U10	DRL ≫ U10**	***.61*** [.28, .94]	DRL ≫ U10**	***.63*** [.56, .69]
Unpracticed subtraction	SUB ≫ U10	SUB ≫ U10***	***.91*** [.69, 1.13]	SUB ≫ U10**	***.69*** [.66, .73]
(Subtraction not	SUB ≫ DRL	SUB ≫ DRL**	***.81*** [.61, 1.02]	SUB ≫ DRL**	***.67*** [.63, .70]
practiced by any condition)	DRL ≈ U10	DRL ≈ U10ns	.19 [.05, .40]	DRL ≈ U10ns	.00 [−.03, .03]
Practiced add-with-8 or -9	U10 ≫ SUB	U10 ≫ SUB***	***.87*** [.62, 1.12]	U10 ≫ SUB**	***1.09*** [1.04, 1.13]
(Add 8 or 9 practiced	U10 > DRL	U10 ≈ DRLns	−.11 [−.41, .19]	U10 ≈ DRLns	−.13 [−.19, −.07]
by U10 and DRL conditions)	DRL > SUB	DRL ≫ SUB***	***1.04*** [.79, 1.28]	DRL ≫ SUB***	***1.35*** [1.31, 1.40]
Unpracticed add-with-8 or	U10 ≫ SUB	U10 > SUB++	***.44*** [.21, .67]	U10 ≫ SUB*	***.57*** [.53, .61]
-9 (add 8 or 9 not	U10 ≫ DRL	U10 > DRLns	***.25*** [.05, .40]	U10 > DRLns	***.28*** [.23, .33]
practiced by any condition)	DRL ≈ SUB	DRL ≈ SUBns	.23 [.00, .52]	DRL ≈ SUBns	***.37*** [.34, .40]

Note. F-index = fluency index; SUB = subtraction intervention; DRL = drill (practice-only) intervention; U10 = Use-10 intervention; ≫ = high relative gain; > = moderate relative gain; ≈ = similar gain. Hedges' g in bold italics meets guidelines for effective practice.
$^{+}$ $p < .10$. $^{++}$ $p < .05$ but comparison is only marginally significant after Benjamini-Hochberg correction. ** $p < .01$. *** $p < .001$.

H1: The Efficacy of the Subtraction Intervention

Mixed-model ANCOVA. Analyses did not reveal a statistically significant main effect of the within participants variable 'experience with items' (i.e., students across all groups did not perform differently on items they practiced compared to item they did not practice), $F(1, 75) = .64$, $p = .426$, but did reveal a statistically significant effect of the between participants variable 'condition,' $F(2, 74) = 15.69$, $p < .001$.

Practiced subtraction items. For subtraction combinations practiced by the subtraction and drill groups, both groups statistically outperformed the use-10 (active-control) group, $F(1, 75) = 25.28$, $p < .001$, $g = 1.00$ (95% CI = 0.69 to 1.31), and $F(1, 75) = 9.39$, $p = .003$, $g = 0.61$ (95% CI = 0.28 to 0.94), respectively.

Unpracticed subtraction items. As predicted for subtraction items not practiced by any group, the subtraction group statistically outperformed both the use-10 (active-control) group, $F(1, 75) = 17.10$, $p < .001$, $g = 0.91$ (95% CI = 0.69 to 1.13), and the drill group, $F(1, 75) = 11.05$, $p = .001$, $g = 0.81$ (95% CI = 0.61 to 1.02). The drill group did not outperform the use-10 group, $F(1, 75) = 0.54$, $p = .464$, $g = 0.19$ (95% CI = 0.05 to 0.40). See the online supplemental materials for additional evidence regarding transfer.

H2: The Efficacy of the Use-10 Intervention

Mixed-model ANCOVA. Analyses did not reveal a statistically significant main effect of the within participants variable 'experience with items,' $F(1, 75) = 2.38$, $p = .127$. However, there was a statistically significant effect of 'condition,' $F(2, 74) = 5.98$, $p = .001$ and, more importantly, a statistically significant experience x condition interaction, $F(2, 74) = 7.97$, $p = .001$.

Practiced add-with-8 or -9 items. For add-with-8 or -9 items practiced by the use-10 and drill groups, both groups statistically outperformed the subtraction (active-control) group, $F(1, 75) = 20.01$, $p < .001$, $g = 0.87$ (95% CI = 0.62 to 1.12), and $F(1, 75) = 27.00$, $p < .001$, $g = 1.04$ (95% CI = 0.79 to 1.28), respectively. The performance of the use-10 and drill groups did not differ statistically, $F(1, 75) = 0.41$, $p = .524$, $g = -0.11$ (95% CI = −0.41 to 0.19).

Unpracticed add-with-8 or -9 items. For add-with-8 or -9 items not practiced by any group, the use-10 group significantly outperformed the subtraction (active-control) group, $F(1, 75) = 5.78$, $p = .037$, $g = 0.44$ (95% CI = 0.21 to 0.67). However, after the application of the Benjamini-Hochberg correction, this comparison was not statistically significant. The drill group did not significantly outperform the subtraction group $F(1, 75) = 0.64$, $p = .425$, $g = 0.23$ (95% CI = 0.00 to 0.52). Additionally, there were no statistically significant differences between the use-10 and the drill groups, $F(1, 75) = 1.77$, $p = .188$, $g = 0.25$ (95% CI = 0.05 to 0.40).

Discussion

Support for the Hypotheses

H1. The efficacy of the subtraction program was supported. Most importantly, at posttest, the subtraction group exhibited significantly and, as indicated by effect size, substantially greater transfer to unpracticed subtraction items than either comparison group. This was true for both the F-index, which indicated overall progress toward more efficient mental subtraction, and fluency rate. At pretest, the mean F-Index score for the subtraction group on the unpracticed subtraction combinations was 0.56 out of a possible 5, indicating that these participants often responded incorrectly and were scored 0 and, when they responded correctly, typically used a very slow, undetermined strategy or counted and were scored 1. Participants in the comparison conditions were only somewhat more successful at pretest. At posttest, the subtraction group overall demonstrated a significant gain of + 0.88, whereas the comparison groups showed essentially no overall gain in progress toward fluency. The average posttest F-index score of the subtraction group is consistent with typically responding correctly, scoring at least 1, and periodically using a highly deliberate subtraction-as-addition reasoning strategy, scoring 2. Although a score of 2 may also indicate a somewhat slow, covert counting strategy, this alternative explanation is unlikely for two reasons. One is that most counting strategies are difficult to implement with subtrahends of 5 to 8 and, hence, would likely have been error prone (scored as 0) or taken more than 6 seconds (scored as 1; Baroody, 1984). See the online supplemental materials for a detailed analysis of counting-based subtraction strategies. A second reason is that the use of any counting-based subtraction strategy with subtrahends of 5 to 8 would probably have been evident to testers and scored as counting or as 1 point.

Furthermore, participants were fluent on only 5% to 7% of unpracticed subtraction combinations at pretest. Whereas the subtraction group's fluency improved 11% overall at posttest, the use-10 group, which received only regular classroom subtraction instruction, and the drill control, which received regular subtraction instruction plus extra subtraction practice, exhibited negligible improvement in fluency of 2% and 1%, respectively. The fact that the opposite was true for unpracticed add-with-8 or -9 items indicates that the transfer to unpracticed subtraction items by the subtraction groups was due to their experimental intervention and not general improvement in mental arithmetic ability.

Two additional results detailed in the online supplemental materials support the efficacy of the subtraction program. Unlike their counterparts in the comparison groups, most individual participants in the subtraction group made more than minimal gains in overall progress toward fluency and fluency itself on unpracticed subtraction combinations—a result that indicates the subtraction program had a broad impact. At posttest, twice as many participants in the subtraction condition overtly or explicitly used a subtraction-as-addition strategy as did participants in the comparison conditions.

In conclusion, whereas most participants in the subtraction group achieved Phase 2 and some even achieved a small degree of Phase 3 proficiency with unpracticed subtraction combinations, most participants in the comparison groups remained at Phase 1 over the course of a school semester. The transfer results, then, indicate that the subtraction intervention was more efficacious in promoting the learning and even the occasional efficient use of a general reasoning strategy for using known sums to determine differences than typical classroom instruction or typical instruction plus unstructured supplemental practice.

H2. The results basically supported the efficacy of the guided use-10 intervention. Most importantly, at posttest, the use-10 group exhibited, as indicated by effect size, substantively greater transfer to unpracticed add-with-8 or -9 combinations than its active control, the subtraction group, or the drill group on both the F-index and fluency rate measures. The fact that the use-10 group showed relatively little gain with unpracticed subtraction items indicates that their improvement with unpracticed add-with-8 or -9 items was due specifically to the use-10 intervention, not general improvement in mental arithmetic. In brief, unlike the regular classroom instruction received by the active control group or such instruction plus unstructured supplemental practice received by the drill group, the highly guided use-10 intervention was successful in promoting the use of a deliberate, or even fluent, general reasoning strategy and, thus, transfer.

Qualifications and Limitations

Although the findings are promising, nine limitations must be noted.

1. Participants had been screened for developmental readiness and were reasonably fluent with the most basic sums: add-with-0 or −1 and the doubles. Plausibly, the indirect-relations guided programs would not have been successful with less ready students. For example, students appear to discover and use the subtraction-as-addition strategy first in cases where sums are well known such as the doubles (Baroody, Ginsburg, & Waxman, 1983). Discovery of the complementary relation between addition and subtraction without such prerequisite knowledge is less likely.

2. The indirect-relations interventions involved instruction on key prerequisite knowledge for transfer. That is, the subtraction group practiced the addition complements for the unpracticed subtraction items, and the use-10 group received instruction on the add-with-10 items related to the unpracticed add-with-8 or -9 combinations. Transfer would probably have been less substantial without instruction on such prerequisite knowledge for applying the subtraction-as-addition and use-10 strategies.

3. It is not clear which relation or activity or combination of relations or activities in the subtraction and use-10 programs is/are the most effective for the meaningful learning and fluent application of their respective reasoning strategy.

4. Most participants in this study had at least one risk factor for later mathematics difficulties. It is not clear if these effects would generalize to other categories of children such as students who were not at risk for later mathematics difficulties or students with significant mathematics disabilities.

5. This study was designed to evaluate only two specific reasoning strategies and, thus, may not generalize to the other combination families that differ in salience (e.g., the complexity of the relations required) and the prior knowledge required to induce or learn the relation or implement the reasoning strategy fluently (see, e.g., Baroody et al., 2014).

6. The posttest occurred two weeks or somewhat longer after the interventions and, thus, retention of gains over longer periods of time is uncertain.

7. The interventions were administered in a relatively structured research environment—providing evidence of their efficacy. However, evidence of the effectiveness of these programs in less structured classroom or home environments is needed.

8. The fact that use-10 and subtraction interventions did not produce fluency rates of more than 32% and 17% on practiced combinations and unpracticed items, respectively, indicates that amount of practice provided by these interventions was not sufficient for many participants to automatize their targeted reasoning strategy. Research is needed to determine the amount of practice required by students of different ability levels to achieve comprehensive fluency with basic sums and differences and whether indirect-relations intervention programs require less practice than other forms of instruction to achieve such fluency.

9. A larger sample would have improved the balance of participant assignment among the intervention conditions, and the increased power would improve the chances of finding statistically significant results.

Conclusions

Unlike previous intervention efforts with no or an undeveloped conceptual basis (Baroody, 1999, 2013; Baroody et al., 2009, 2011a; 2011b; Murata, 2004; Walker et al., 2013), the present indirect-relations programs produced substantial transfer, as gauged by effect size. These results support the instructional approach to basic mental addition and subtraction recommended by CCSSO (2010) in the *Common Core State Standards* of focusing on relations to promote reasoning strategies and indicate that an explicit model of the meaningful memorization of basic combinations may need to be developed.

Educational Implications

Over the last 100 years, research has prompted a gradual shift from a direct-facts approach to promoting fluency with basic sums and differences toward indirect approaches. Thorndike's (1922) direct-facts approach was the dominant instructional strategy for the first three-quarters of the 20th century or so (Brownell, 1935; Ginsburg, Klein, & Starkey, 1998). Instruction routinely involved bypassing Phases 1 and 2, suppressing students' informal use of counting and reasoning strategies, and using extensive drill to promote memorization of basic facts by rote. As Elliott's (2014) editorial illustrates, some parents, educators, and researchers (Hasselbring et al., 1988; Walker et al., 2013, 2014) still adhere to this approach. As research revealed the utility of children's informal counting-based strategies, textbooks and instructional practice until the 1990s commonly provided Phase 1 experiences—if only briefly—before relying heavily on drill to promote automatic recall or Phase 3 (Thornton, 1978). As research revealed the usefulness of students' informal reasoning strategies, mathematics educators (Rathmell, 1978; Steinberg, 1985; Suydam & Weaver, 1975; Thornton, 1978; Thornton & Toohey, 1985) recommended what has become the popular view and practice of using Phase 2 as a steppingstone from Phase 1 to Phase 3 (CCSSO, 2010; NCTM, 2000; NRC, 2001).

The results of the present RCT provide rare, methodologically sound support for an indirect approach that focuses on the meaningful learning of reasoning strategies for relatively difficult to learn subtraction and add-with-8 or -9 combinations. Consistent with Henry and Brown's (2008) correlational data, the present results indicate that an indirect approach was more fruitful than a direct-facts approach (the drill intervention involving haphazardly ordered drill) in fostering overall progress toward fluency, as measured by the F-index, and fluency itself, as measured by fluency rate, with unpracticed combinations. Contrary to the view that an indirect approach may be too difficult for primary-grade students or is a relatively inefficient use of instructional time, appreciable gains in fluency with practiced combinations were achieved with minimal practice—only 34 repetitions for each practiced-subtraction item and 26 repetitions for each addition complements. Indeed, with the equivalent amount of practice provided by the direct-facts approach of the drill intervention, the indirect-relations approach produced similar fluency retention with practiced subtraction items with the bonus of significantly greater fluency transfer to unpracticed combinations. An indirect-relations approach that focuses on meaningful memorization, which may make instruction and practice more engaging by highlighting relations and encouraging the discovery of reasoning strategies, actually appears to be more efficient than a direct-facts approach geared to memorization of basic facts by rote via unstructured practice (see also Carpenter et al., 1989).

The present results further underscore that an indirect approach needs to be conceptually based, carefully designed, and sustained in order to promote substantial transfer. The evidence of substantial transfer was obtained despite the fact that the regular classroom instruction received by the comparison groups, except for that of the seven Grade K participants, involved lessons on reasoning strategies for subtraction and add-with-9 combinations (again see online supplemental materials for a summary of regular classroom instruction). Given their relatively spotty, superficial, brief, and perhaps overly directive treatment of reasoning strategies (Baroody, 2016), current curricula typically fall short of an effective indirect-relations approach and resemble more closely an indirect-facts approach in which learning numerical relations and reasoning strategies is merely a secondary means to the primary end of fact recall.

The present subtraction intervention, for example, focused on the following experiences and concepts that may be needed to construct a deep understanding of the relations between addition and subtraction and to provide a solid foundation for the meaningful learning of the subtraction-as-addition strategy: *empirical inversion* (e.g., adding 2 items to a collection of 3 and then taking away 2 items), the *undoing concept* (adding and then subtraction the same amount or number [or vice versa] undo each other), and the *shared parts and whole concept* (recognizing that complementary addition and subtraction items have in common the same parts and the same whole). See Baroody (2016) or the online supplemental materials for a hypothetical learning progression.

In contrast, a review of one supplementary curriculum and six mainstream Grade-1 curricula—including the two used in the participating classrooms that produced a relatively small impact on progress toward subtraction fluency or fluency itself by the comparison groups—indicated an incomplete, superficial, nonsystematic, and brief treatment of the possible foundations for the meaningfully learning the subtraction-as-addition strategy (Baroody, 2016). Only the supplementary curricula directly, explicitly, and consistently capitalized on empirical inversion and the undoing concept. Moreover, although all six mainstream curricula included instruction on the idea that complementary addition and subtraction combinations involve the same three numbers (shared numbers concept), none regularly and explicitly related this concept to part-whole relations (e.g., $5 + 3 = 8$ and $8 - 3 = 5$ have the three same numbers: the whole 8 and the parts 3 and 5). Understanding the undoing concept and part-whole relations may help children understand why addition and subtraction complements share the same three numbers and the complement principle itself (e.g., if the part 5 and the part 3 make the whole 8, then subtracting one of the parts from the whole 8 leaves the other part). Only three of the seven curricula surveyed allude to the complement principle, the rationale for the subtraction-as-addition strategy. However, none of these curricula describe the principle meaningfully in terms of part-whole relations (i.e., provide a clear explanation of why a known sum can be used to determine an unknown difference). Unfortunately, the barren references to the complement principle used (e.g., if "$3 + 5 = 8$, then 8 - 5 = what?") may not help many students, particularly those struggling with mathematics, to understand the if-then connection between addition and subtraction and use it meaningfully to determine unknown differences. In brief, taken together, the foregoing analysis and the results of the present study suggest that existing curricula and teaching practices regarding the promotion of basic combination fluency need to be carefully analyzed and refined to more effectively embody the indirect-relations approach.

Finally, the present results add to the growing evidence that digital environments involving games can provide engaging supplemental learning opportunities in mathematics education (Chang, Wu, Weng, & Sung, 2012; Ke, 2013; Kebritchi, Hirumi, & Bai, 2010; Obersteiner, Reiss, & Ufer, 2013; Pareto, Arvemo, Dahl, Haake, & Gulz, 2011; Pea, 1987; Shin, Sutherland, Norris, & Soloway, 2012). Supplemental instructional software may be particularly useful in implementing a relatively sophisticated indirect-relations approach for relatively difficult basic add-with-8 or -9 and subtraction combinations.

Theoretical Implications

The present results indicate the need for a new and more powerful model of mental arithmetic development. For example, neither the distribution-of-associations model (Shrager & Siegler, 1998; Siegler, 1987; Siegler & Araya, 2005) nor the identical-elements model (Rickard, 2005; Walker et al., 2013, 2014) adequately accounts for how relational learning affects the memorization of basic sums or differences or for the transfer of fluency. The evidence that practice with the addition complements did translate into fluency with unpracticed subtraction combinations supports the complement problem mediation hypothesis. This evidence and the significant transfer of fluency to unpracticed add-with-8 or -9 items suggests that both the distribution-of-associations model and the identical-elements model may be a reasonable representation of the process of memorizing the basic subtraction facts by rote but not meaningful memorization of basic sums and differences (cf. Bisanz, 2003).

The indirect-facts model suggested by National Mathematics Advisory Panel (NMAP, 2008) is a step toward a more powerful model in that it includes aspects of both the direct-facts and indirect-relations approaches. Specifically, the model underscores the importance of repeated practice and the value of structuring practice to facilitate the learning of arithmetic relations. Consistent with the evidence of fluency transfer in the present study, the model further suggests that the retrieval system may be composed of both fact recall and reasoning processes (Campbell & Beech, 2014; Fayol & Thevenot, 2012; NMAP, 2008; Verschaffel et al., 2007). Nevertheless, existing indirect-facts models do not explicitly account for how relatively meager practice can produce reasoning-based transfer at least to a modest degree.

Moreover, numerous issues remain unresolved and in need of future research. Might the retrieval network that is the product of the meaningful memorization of basic combinations be organized and function differently than a network that is product of memorization by rote? How exactly might the development of conceptual knowledge facilitate the learning of related combinations and affect the representation of basic sums and difference? Are conceptual, procedural, and factual knowledge stored and processed independently; semi-interdependently with separate representations that interact cooperatively during retrieval of sums and differences; or interdependently in a single knowledge web that embodies facts, relations, and operations (Baroody & Varma, 2006)? Are relational and factual knowledge accessed and applied simultaneously in parallel or sequentially (Baroody & Varma, 2006)?

Consider, for example, the possible different scenarios for how learned relations between addition and subtraction and the subtraction-as-addition strategy might be represented in long-term memory (LTM) and affect the retrieval of basic differences. With modular models, relational knowledge of addition and subtraction, Phase 1 and 2 strategies, and factual knowledge of subtraction combinations are stored in separate sites within LTM and processed independently. Phase 1 and 2 experiences may facilitate the learning of the association between a subtraction expression and its difference, but conceptual or procedural knowledge does not affect the automatic fact recall process once initiated. According to classic versions of this noncooperative modular view, the retrieval system is composed exclusively of the factual knowledge, and experts rely exclusively on automatic fact recall, a reproductive process. According to sequential processing versions, the efficient fact recall is activated first, and the subtraction-as-addition strategy, which is a less efficient reconstructive process, is activated as a backup if fact recall fails. If the subtraction-as-addition strategy has not been learned or is relatively inefficient, another reconstructive process, namely a counting strategy, is deployed. According to simultaneous processing versions, reconstructive and reproductive processes compete in a "horse race," and the more efficient process generates the difference. With both classic versions of the modular view, complement problem mediation in the form of transfer of a practice effect would not be evident in experts' factual recall but would be evident in the case of nonexperts' use of a backup strategy.

Alternatively, according to the modular model suggested by NMAP (2008), the retrieval system could be composed of compiled or automatic reconstructive strategies as well as a reproductive strategy. Although NMAP (2008) concluded "mediated retrieval" is slower and more error-prone than fact recall, automatic reconstructive processes may, in fact, be as efficient, or even more efficient than, a reproductive process, which may be subject to associative interference or confusion (Baroody, 1994; Baroody & Varma, 2006; Barrouillet & Thevenot, 2013; Campbell & Beech, 2014; Fayol & Thevenot, 2012). In either case, the type of retrieval process selected first or the winner of the horse race would depend on an individual's learning history, and complement problem mediation in the form of transfer of a practice effect may or may not be evident in her fluent retrieval.

One possible partially modular model is that retrieval involves a sequential process of first activating conceptual knowledge of the complementary relation between addition and subtraction or compiled procedural knowledge of the subtraction-as-addition strategy and then initiating a search elsewhere in LTM of the factual knowledge of the related addition combination. With an even more extreme nonmodular model, learning relations between addition and subtraction might prompt a wholesale reorganization of retrieval system in which sums and differences are represented as an integrated "mental fact triangle" (Baroody, 1985). Retrieval in such a system would be essentially simultaneous processing of integrated relational and factual knowledge. The partially modular or the nonmodular model would be applicable only in cases of meaningful memorization. So again, complement problem mediation in the form of transfer of a practice effect should be evident in the fluent retrieval of an expert who memorized basic subtraction combinations meaningfully but not of one who memorized such combinations by rote.

Complicating matters is the fact that different families of subtraction combinations may be meaningfully learned and represented differently. Whereas the more difficult subtraction items studied in the present research may involve learning and compiling the addition-as-subtraction strategy, other classes of subtraction combinations may entail other reconstructive processes. For example, retrieval involving the subtraction of 0 or 1 might involve the $n - 0 = n$ or subtractive identity rule and the number-before n rule (e.g., the difference of $7 - 1$ is the number before n in the counting sequence). Subtraction doubles such as $7 - 7$ might involve the $n - n = 0$ or subtractive negation rule. The subtraction of number neighbors such as $8 - 7$ might entail the difference of 1 rule: "the difference of two adjacent numbers in the counting sequence is always 1." In brief, although mental arithmetic is one of the most thoroughly studied domains in educational psychology (NMAP, 2008), much remains to be resolved about the meaningful memorization of basic sums and differences and the resulting retrieval network. Such a resolution probably cannot be achieved without taking into account a participant's learning history and, within participants, the possibility that a learner, or even an expert, is using different strategies for different classes of combinations.

References

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44,* 75–106. http://dx.doi.org/10.1016/0010-0277(92)90051-I

Ashcraft, M. H., & Guillaume, M. M. (2009). Mathematical cognition and the problem size effect. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 51, pp. 121–151). Burlington, VT: Academic Press. http://dx.doi.org/10.1016/S0079-7421(09)51004-3

Baroody, A. J. (1984). Children's difficulties in subtraction: Some causes and questions. *Journal for Research in Mathematics Education, 15,* 203–213. http://dx.doi.org/10.2307/748349

Baroody, A. J. (1985). Mastery of the basic number combinations: Internalization of relationships or facts? *Journal for Research in Mathematics Education, 16,* 83–98. http://dx.doi.org/10.2307/748366

Baroody, A. J. (1994). An evaluation of evidence supporting fact-retrieval models. *Learning and Individual Differences, 6,* 1–36. http://dx.doi.org/10.1016/1041-6080(94)90013-2

Baroody, A. J. (1999). Children's relational knowledge of addition and subtraction. *Cognition and Instruction, 17,* 137–175. http://dx.doi.org/10.1207/S1532690XCI170201

Baroody, A. J. (2013). *Final Report for Institute of Education Sciences Grant R305A080479*. Champaign, IL: University of Illinois.

Baroody, A. J. (2016). Curricular approaches to introducing subtraction and fostering fluency with basic differences in grade 1. *Monograph of the Journal of Pensamiento Numérico y Algebraico, 10*.

Baroody, A. J., Bajwa, N. P., & Eiland, M. (2009). Why can't Johnny remember the basic facts? *Developmental Disabilities Research Reviews, 15,* 69–79. http://dx.doi.org/10.1002/ddrr.45

Baroody, A. J., Eiland, M. D., Bajwa, N. P., & Baroody, S. C. (2009, June). *Fostering primary-level at-risk children's fluency with basic addition combinations*. Poster presented at the Fourth Annual IES Research Conference, Washington, DC.

Baroody, A. J., Eiland, M. D., & Baroody, S. C. (2011a, April). *Fostering first-graders' fluency with basic addition combinations*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Baroody, A. J., Eiland, M. D., & Baroody, S. C. (2011b, April). *Fostering third-graders' fluency with basic addition and subtraction combinations*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Baroody, A. J., Eiland, M. D., Purpura, D. J., & Reid, E. E. (2012). Fostering kindergarten children's number sense. *Cognition and Instruction, 30,* 435–470. http://dx.doi.org/10.1080/07370008.2012.720152

Baroody, A. J., Eiland, M. D., Purpura, D. J., & Reid, E. E. (2013). Can computer-assisted discovery learning foster first graders' fluency with the most basic addition combinations? *American Educational Research Journal, 50,* 533–573. http://dx.doi.org/10.3102/0002831212473349

Baroody, A. J., Ginsburg, H. P., & Waxman, B. (1983). Children's use of mathematical structure. *Journal for Research in Mathematics Education, 14,* 156–168. http://dx.doi.org/10.2307/748379

Baroody, A. J., Purpura, D. J., Eiland, M. D., & Reid, E. E. (2014). Fostering first-graders' fluency with basic subtraction and larger addition combinations via computer-assisted instruction. *Cognition and Instruction, 32,* 159–197. http://dx.doi.org/10.1080/07370008.2014.887084

Baroody, A. J., Purpura, D. J., Eiland, M. D., & Reid, E. E. (2015). The impact of highly and minimally guided discovery instruction on promoting the learning of reasoning strategies for basic add-1 and doubles combinations. *Early Childhood Research Quarterly, 30,* 93–105. http://dx.doi.org/10.1016/j.ecresq.2014.09.003

Baroody, A. J., & Tiilikainen, S. H. (2003). Two perspectives on addition development. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 75–125). Mahwah, NJ: Lawrence Erlbaum.

Baroody, A. J., & Varma, S. (2006, November). *The active construction view of basic number fact knowledge: New directions for cognitive neuroscience*. Invited presentation at the Neural Basis of Mathematical Cognition Conference, Nashville, TN.

Barrouillet, P., Mignon, M., & Thevenot, C. (2008). Strategies in subtraction problem solving in children. *Journal of Experimental Child Psychology, 99,* 233–251. http://dx.doi.org/10.1016/j.jecp.2007.12.001

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. http://dx.doi.org/10.1016/j.cognition.2013.02.018

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B. Methodological, 57,* 289–300.

Bezuk, N. S., & Cegelka, P. T. (1995). Effective mathematics instruction for all students. In P. T. Cegelka & W. H. Berdine (Eds.), *Effective instruction for students with learning difficulties* (pp. 345–384). Needham Heights, MA: Allyn & Bacon.

Bisanz, J. (2003). Arithmetical development: Commentary on chapters 1 through 8 and reflections on directions. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 435–452). Mahwah, NJ: Lawrence Erlbaum.

Boaler, J. (2009). *What's math got to do with it? How parents and teachers can help children learn to love their least favorite subject*. New York, NY: Penguin.

Boaler, J. (2015). *Fluency without fear: Research evidence on the best ways to learn math facts*. Retrieved from http://www.youcubed.org/fluency-without-fear/

Brownell, W. A. (1935). Psychological considerations in the learning and the teaching of arithmetic. In W. D. Reeve (Ed.), *The teaching of arithmetic* (Tenth yearbook, National Council of Teachers of Mathematics, pp. 1–31). New York, NY: Teachers College, Columbia University.

Brownell, W. A. (1941). *Arithmetic in grades I and II: A critical summary of new and previously reported research*. Durham, NC: Duke University Press.

Campbell, J. I. D. (1987). The role of associative interference in learning and retrieving arithmetic facts. In J. A. Sloboda & D. Rogers (Eds.), *Cognitive process in mathematics: Keele cognition seminars* (Vol. 1, pp. 107–122). New York, NY: Clarendon Press Press.

Campbell, J. I. D. (2008). Subtraction by addition. *Memory & Cognition, 36,* 1094–1102. http://dx.doi.org/10.3758/MC.36.6.1094

Campbell, J. I. D., & Agnew, H. (2009). Retrieval savings with nonidentical elements: The case of simple addition and subtraction. *Psychonomic Bulletin & Review, 16,* 938–944. http://dx.doi.org/10.3758/PBR.16.5.938

Campbell, J. I. D., & Beech, L. C. (2014). No generalization of practice for nonzero simple addition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 1766–1771. http://dx.doi.org/10.1037/xlm0000003

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. http://dx.doi.org/10.3758/BF03193585

Campbell, J. I. D., & Graham, D. J. (1985). Mental multiplication skill: Structure, process, and acquisition. *Canadian Journal of Psychology, 39,* 338–362. http://dx.doi.org/10.1037/h0080065

Campbell, J. I. D., & Therriault, N. H. (2013). Retrieval-induced forgetting of arithmetic fact but not rules. *Journal of Cognitive Psychology, 25,* 717–724. http://dx.doi.org/10.1080/20445911.2013.798328

Canobi, K. H. (2009). Concept-procedure interactions in children's addition and subtraction. *Journal of Experimental Child Psychology, 102,* 131–149. http://dx.doi.org/10.1016/j.jecp.2008.07.008

Canobi, K. H., Reeve, R. A., & Pattison, P. E. (1998). The role of conceptual understanding in children's addition problem solving. *Developmental Psychology, 34,* 882–891. http://dx.doi.org/10.1037/0012-1649.34.5.882

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C. P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal, 26,* 499–532. http://dx.doi.org/10.3102/00028312026004499

Chang, K. E., Wu, L. J., Weng, S. E., & Sung, Y. T. (2012). Embedding game-based problem-solving phase into problem-posing system for mathematics learning. *Computers & Education, 58,* 775–786. http://dx.doi.org/10.1016/j.compedu.2011.10.002

Council of Chief State School Officers. (2010). *Common Core State Standards: Preparing America's students for college and career*. Retrieved from http://www.corestandards.org/

Cowan, R. (2003). Does it all add up? Changes in children's knowledge of addition combinations, strategies, and principles. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 35–74). Mahwah, NJ: Lawrence Erlbaum.

Crowley, K., Shrager, J., & Siegler, R. S. (1997). Strategy discovery as a competitive negotiation between metacognitive and associative mechanisms. *Developmental Review, 17,* 462–489. http://dx.doi.org/10.1006/drev.1997.0442

Dowker, A. (2009). Use of derived fact strategies by children with mathematical difficulties. *Cognitive Development, 24,* 401–410. http://dx.doi.org/10.1016/j.cogdev.2009.09.005

Eiland, M. D. (2014). *Fostering fluency with basic addition and subtraction facts using computer-aided instruction*. [Doctoral dissertation]. Retrieved from http://hdl.handle.net/2142/50759

Elliott, J. (2014). Guest commentary: Parents would like answers on curriculum. *Mahomet Citizen, 146, 4*, 9.

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

Fuson, K. C. (2006). *Math expressions*. Boston, MA: Houghton Mifflin.

Ginsburg, H. P., & Baroody, A. J. (2003). *Test of early mathematics ability* (3rd ed.). Austin, TX: Pro-Ed.

Ginsburg, H. P., Klein, A., & Starkey, P. (1998). The development of children's mathematical knowledge: Connecting research with practice. In I. E. Sigel & K. A. Renninger (Eds.), *Handbook of child psychology: Vol. 4. Child psychology in practice* (5th ed., pp. 401–476). New York, NY: Wiley.

Goldman, S. R., & Pellegrino, J. (1986). Microcomputer: Effective drill and practice. *Academic Therapy, 22,* 133–140.

Hasselbring, T. S., Goin, L. T., & Bransford, J. D. (1988). Developing math automaticity in learning handicapped children: The role of computerized drill and practice. *Focus on Exceptional Children, 20,* 1–7.

Henry, V., & Brown, R. (2008). First-grade basic facts: An investigation into teaching and learning of an accelerated, high demand memorization standard. *Journal for Research in Mathematics Education, 39,* 153–183.

Institute of Education Sciences. (2014). *What works clearinghouse: Procedures and standards handbook (Version 3)*. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_standards_handbook.pdf

Jordan, N. C. (2007). The need for number sense. *Educational Leadership, 65,* 63–66.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2013). The cost of concreteness: The effect of nonessential information on analogical transfer. *Journal of Experimental Psychology, 19,* 14–29. http://dx.doi.org/10.1037/a0031931

Ke, F. (2013). Computer-game-based tutoring of mathematics. *Computers & Education, 60,* 448–457. http://dx.doi.org/10.1016/j.compedu.2012.08.012

Kebritchi, M., Hirumi, A., & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers & Education, 55,* 427–443. http://dx.doi.org/10.1016/j.compedu.2010.02.007

Koscinski, S. T., & Gast, D. L. (1993a). Use of constant time delay in teaching multiplication facts to students with learning disabilities. *Journal of Learning Disabilities, 26,* 533–544, 567. http://dx.doi.org/10.1177/002221949302600807

Koscinski, S. T., & Gast, D. K. (1993b). Computer-assisted instruction with constant time delay to teach multiplication facts to students with learning disabilities. *Learning Disabilities Research & Practice, 8,* 157–168.

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., . . . Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms*. Washington, DC: IES National Center for Special Education Research, Institute of Education Sciences.

Logan, G. D. (1991). Automaticity and memory. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human* *memory in honor of Bennet B. Murdock* (pp. 347–366). Hillsdale, NJ: Lawrence Erlbaum.

Murata, A. (2004). Paths to learning ten-structured understandings of teen sums: Addition solution methods of Japanese grade 1 students. *Cognition and Instruction, 22,* 185–218. http://dx.doi.org/10.1207/s1532690xci2202_2

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics: Standards 2000*. Reston, VA: Author.

National Council of Teachers of Mathematics. (2006). *Curriculum focal points for prekindergarten through grade 8 mathematics*. Reston, VA: Author.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.

National Research Council. (2001). *Adding it up: Helping children learn mathematics* (J. Kilpatrick, J. Swafford, & B. Findell, Eds.). Washington, DC: National Academy Press.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Obersteiner, A., Reiss, K., & Ufer, S. (2013). How training on exact or approximate mental representations of number can enhance first-grade students' basic number processing and arithmetic skills. *Learning and Instruction, 23,* 125–135. http://dx.doi.org/10.1016/j.learninstruc.2012.08.004

Pareto, L., Arvemo, T., Dahl, Y., Haake, M., & Gulz, A. (2011). A teachable-agent arithmetic game's effects on mathematics understanding, attitude and self-efficacy. In *Artificial intelligence in education* (pp. 247–255). http://dx.doi.org/10.1007/978-3-642-21869-9_33

Pea, R. D. (1987). Cognitive technologies for mathematics education. In A. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 89–122). Hillsdale, NJ: Lawrence Erlbaum; Retrieved from https://telearn.archives-ouvertes.fr/hal-00190547

Peters, G., De Smedt, B., Torbeyns, J., Ghesquière, P., & Verschaffel, L. (2010). Adults' use of subtraction by addition. *Acta Psychologica, 135,* 323–329. http://dx.doi.org/10.1016/j.actpsy.2010.08.007

Putnam, R. T., deBettencourt, L. U., & Leinhardt, G. (1990). Understanding of derived fact strategies in addition and subtraction. *Cognition and Instruction, 7,* 245–285. http://dx.doi.org/10.1207/s1532690xci0703_3

Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development, 14,* 187–202. http://dx.doi.org/10.1080/15248372.2012.664593

Rathmell, E. C. (1978). Using thinking strategies to teach basic facts. In M. N. Suydam & R. E. Reys (Eds.), *Developing computational skills* (1978 Yearbook, pp. 13–50). Reston, VA: National Council of Teachers of Mathematics.

Rickard, T. C. (2005). A revised identical elements model of arithmetic fact representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 250–257. http://dx.doi.org/10.1037/0278-7393.31.2.250

Robinson, K. M., & Dubé, A. K. (2009). Children's understanding of addition and subtraction concepts. *Journal of Experimental Child Psychology, 103,* 532–545. http://dx.doi.org/10.1016/j.jecp.2008.12.002

Robinson, K. M., Ninowski, J. E., & Gray, M. L. (2006). Children's understanding of the arithmetic concepts of inversion and associativity. *Journal of Experimental Child Psychology, 94,* 349–362. http://dx.doi.org/10.1016/j.jecp.2006.03.004

Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research: Learning trajectories for young children*. New York, NY: Routledge.

Shin, N., Sutherland, L. M., Norris, C. A., & Soloway, E. (2012). Effects of game technology on elementary student learning in mathematics. *British Journal of Educational Technology, 43,* 540–560. http://dx.doi.org/10.1111/j.1467-8535.2011.01197.x

Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science, 9,* 405–410. http://dx.doi.org/10.1111/1467-9280.00076

Siegler, R. S. (1987). Strategy choices in subtraction. In J. Sloboda & D. Rogers (Eds.), *Cognitive process in mathematics* (pp. 81–106). New York, NY: Oxford University Press.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New-York, NY: New York: Oxford University Press.

Siegler, R., & Araya, R. (2005). A computational model of conscious and unconscious strategy discovery. In R. Kail (Ed.), *Advances in child development and behaviour* (Vol. 33, pp. 1–42). New York, NY: Academic Press. http://dx.doi.org/10.1016/S0065-2407(05)80003-5

Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, NJ: Erlbaum.

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

Silbert, J., Carnine, D., & Stein, M. (1990). *Direct instruction mathematics* (2nd ed.). Columbus, OH: Merrill.

Steinberg, R. M. (1985). Instruction on derived facts strategies in addition and subtraction. *Journal for Research in Mathematics Education, 16,* 337–355. http://dx.doi.org/10.2307/749356

Suydam, M., & Weaver, J. F. (1975). Research on mathematics learning. In J. N. Payne (Ed.), *Mathematics learning in early childhood* (pp. 44–67). Reston, VA: NCTM.

Thorndike, E. L. (1922). *The psychology of arithmetic*. New York, NY: Macmillan. http://dx.doi.org/10.1037/11080-000

Thornton, C. A. (1978). Emphasizing thinking strategies in basic fact instruction. *Journal for Research in Mathematics Education, 9,* 214–227. http://dx.doi.org/10.2307/748999

Thornton, C., & Toohey, M. (1985). Basic math facts: Guidelines for teaching and learning. *Learning Disabilities Focus, 1,* 44–57.

Tobias, S. (1993). *Overcoming math anxiety*. New York, NY: Norton & Co.

Torgesen, J. K., & Young, K. A. (1983). Priorities for the use of microcomputers with learning disabled children. *Journal of Learning Disabilities, 16,* 234–237. http://dx.doi.org/10.1177/002221948301600413

University of Chicago School Mathematics Project. (2005). *Everyday mathematics teacher's lesson guide* (Vol. 1). Columbus, OH: McGraw-Hill.

Verschaffel, L., Greer, B., & De Corte, E. (2007). Whole number concepts and operations. In F. K. Lester, Jr., (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 557–628). Charlotte, NC: NCTM & Information Age Publishing.

Walker, D., Bajic, D., Mickes, L., Kwak, J., & Rickard, T. C. (2014). Specificity of children's arithmetic learning. *Journal of Experimental Child Psychology, 122,* 62–74. http://dx.doi.org/10.1016/j.jecp.2013.11.018

Walker, D., Mickes, L., Bajic, D., Nailon, C. R., & Rickard, T. C. (2013). A test of two methods of arithmetic fluency training and implications for educational practice. *Journal of Applied Research in Memory & Cognition, 2,* 25–32. http://dx.doi.org/10.1016/j.jarmac.2013.02.001

Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54,* 594–604. http://dx.doi.org/10.1037/0003-066X.54.8.594

Received February 25, 2014
Revision received June 5, 2015
Accepted June 5, 2015 ■