*American Educational Research Journal*
*June 2013, Vol. 50, No. 3, pp. 533–573*
*DOI: 10.3102/0002831212473349*


# Can Computer-Assisted Discovery Learning Foster First Graders' Fluency With the Most Basic Addition Combinations?

Arthur J. Baroody
Michael D. Eiland
*University of Illinois at Urbana-Champaign*
David J. Purpura
*Purdue University*
Erin E. Reid
*University of Illinois at Urbana-Champaign*

*In a 9-month training experiment, 64 first graders with a risk factor were randomly assigned to computer-assisted structured discovery of the add-1 rule (e.g., the sum of 7 + 1 is the number after "seven" when we count), unstructured discovery learning of this regularity, or an active-control group. Planned contrasts revealed that the add-1 conditions were more effective than regular instruction/practice in promoting the learning of the add-1 rule. Contrary to the conclusions of Alfieri, Brooks, Aldrich, and Tenenbaum (2011) and Kirschner, Sweller, and Clark (2006), participants in the structured add-1 condition did not outperform those in the unstructured add-1 group on practiced and unpracticed* n *+ 1 and 1 +* n *items at the posttests. The control participants did not exhibit evidence of learning*

Arthur J. Baroody is a professor emeritus of curriculum and instruction, College of Education, University of Illinois at UrbanaChampaign, Champaign, IL 61820; e-mail: *baroody@illinois.edu*. His research focuses on number and arithmetic teaching and learning in early childhood.

Michael D. Eiland is currently a doctoral candidate in early childhood and elementary mathematics education in the Department of Curriculum and Instruction, College of Education, University of Illinois at Urbana–Champaign.

David J. Purpura is an assistant professor in the Department of Human Development and Family Studies at Purdue University. His research focus is on the development of early mathematics skills in young children and how nonmathematical factors influence this development.

Erin E. Reid is currently an IES postdoctoral fellow in the Department of Curriculum and Instruction, College of Education, University of Illinois at Urbana–Champaign. Her research focus is on mathematical development in young children and early intervention.

*a general near-doubles reasoning strategy (if 4 + 4 is 8 and 4 + 5 = 4 + 4 + 1, then the sum of 4 + 5 must be 9). The add-1, but not the active-control, participants achieved success, including transfer, because the former had mastered the developmental prerequisites for add-1 rule and the latter had not mastered the prerequisites for the near-doubles strategy.*

KEYWORDS: elementary math, instructional design, discovery learning, computer-assisted instruction, basic number facts

*Fluency* with basic addition combinations involves appropriately and adaptively, as well as quickly and accurately (efficiently), generating sums to single-digit items such as 7 + 1 and 3 + 4. Appropriate use entails selectively applying knowledge to suitable problems only. For example, the *add-1 rule* specifies that the sum of $n + 1$ or $1 + n$—but not other—items is the number after $n$ in the count sequence (e.g., the sum of 7 + 1 or 1 + 7, but not 7 + 0 or 2 + 7, is the number after seven—eight). Adaptive use involves flexibly adjusting or applying knowledge to solve new problems (e.g., transfer of the add-1 rule to previously unpracticed $n + 1$ or $1 + n$ items, including those where $n$ is a multidigit number). Fluency with basic sums is widely recognized as a critical expertise *all* children need (e.g., National Council of Teachers of Mathematics [NCTM], 2000, 2006; National Mathematics Advisory Panel [NMAP], 2008; National Research Council [NRC], 2001). However, how best to achieve this goal is still not entirely clear. Particularly unclear is what role discovery learning might play in promoting fluency with different combination families and to what extent such instruction needs to be guided. Alfieri, Brooks, Aldrich, and Tenenbaum (2011) defined discovery learning as not providing learners with the target information or conceptual information but creating the opportunity to "find it independently . . . with only the provided materials" (p. 2).

## Rationale for the Present Study

The primary aim of the present study was to evaluate the promise or feasibility of computer programs designed to foster first graders' fluency with the most basic sums. The feasibility study also provided the opportunity to explore the relative value of different forms of discovery learning and the implications of a hypothetical learning progression of mental addition. The background and justifications for the primary aim and two subsidiary aims are discussed in turn.

### Primary Aim: Feasibility of Software for Promoting Fluency With Basic Sums

Reasoning strategies—using known sums and relations to deduce the sum of an unknown item—play a central role in achieving fluency with basic sums. For the near-double 2 + 3, for example, the *near-doubles strategy* entails using the known double (2 + 2 = 4) and the relation that "3 is 1 more than 2" to reason that 2 + 3 must be 5. Discussed, in turn, are why reasoning strategies are a key basis for achieving fluency with basic sums, two of the most basic reasoning strategies, and the reason for developing and evaluating interventions for these two strategies (preventing all too common learning difficulties).

*Reasoning strategies as a key basis for fluency.* The conventional wisdom regarding the development of mental addition has shifted dramatically over the past century (Verschaffel, Greer, & De Corte, 2007). For most of the 20th century, behavioral/cognitive psychologists and educators typically viewed learning the basic sums as the simple process of memorizing facts by rote to achieve *efficient* recall. Practice was viewed as *the* basis for increasing associative strength and promoting fact retrieval (Ashcraft, 1992), or at least the most important factor in this process (Siegler, 1986; Torgeson & Young, 1983). Thorndike's (1922) *law of frequency* (the more two stimuli are presented together, the stronger their association) justified large doses of practice. For example, according to the distribution-of-associations model (Siegler & Jenkins, 1989) and its successors (e.g., Shrager & Siegler, 1998), a memory trace is laid down each time an expression and its sum (e.g., 5 + 1 = 6) is practiced, and thousands of such traces are necessary to achieve efficient fact recall. Moreover, each fact was presumed to accumulate associative strength independently of other, even related, combinations (e.g., practice with 5 + 1 was presumed to have no effect on its commuted partner 1 + 5).

Research that examined how children solve sums revealed that they use informal strategies before memorizing combinations (for a review, see Cowan, 2003). Initially, they use counting strategies involving object or verbal counting (e.g., for 2 + 3, counting "Three, four is 1 more, *five* is two more—the answer is five"). Then, as a result of discovering patterns or relations, children invent reasoning strategies, such as the near-doubles strategy. Early theorists regarded both counting *and* reasoning strategies as crutches or bad habits that interfered with the process of memorizing facts by rote (see e.g., reviews by Baroody, 1985; Cowan, 2003). More recently, such strategies have been considered a valuable source of practice for memorizing sums and connecting such factual knowledge to meaningful concepts (Crowley, Shrager, & Siegler, 1997; Fuchs et al., 2010; NMAP, 2008; NRC, 2001).

There is a growing recognition that reasoning strategies play a wider role in achieving *fluent* (as opposed to simply efficient) mental addition. Besides providing practice to reinforce the association between a particular expression and its sum (which promotes efficient recall), reasoning strategies can serve as a bridge to fluent retrieval in two ways: (a) Such strategies "enable pupils to organize and understand relations among facts that aid in [meaningful] memorization and recall" (Rathmell, 1978, p. 16)—that is, provide them an organizing framework for storing combinations (Canobi, Reeve, & Pattison, 1998; Dowker, 2009; Sarama & Clements, 2009). (b) With practice, deliberate (slow and conscious) reasoning strategies can become automatic (efficient and nonconscious) and serve as a component of the retrieval system—an efficient memory network that is composed of both factual and relational knowledge (Baroody, 1985; Fayol & Thevenot, 2012; Jerman, 1970; NMAP, 2008). Children, then, typically progress through three overlapping phases in the *meaningful learning* of a particular basic sum or family of sums: (a) Phase 1 (counting strategies), (b) Phase 2 (conscious or deliberate reasoning strategies), and (c) Phase 3 (fluent retrieval—including nonconscious or automatic reasoning strategies; Baroody, Bajwa, & Eiland, 2009; Baroody & Varma, in press; Fayol & Thevenot, 2012; Verschaffel et al., 2007). In brief, Phase 2 is a vital bridge between Phases 1 and 3, and the relational learning and the automatic reasoning processes it yields provides an important basis for the appropriate and adaptive—as well as efficient—application of basic addition knowledge.

Indeed, research indicates that meaningful memorizations is more effective than memorization by rote in general (Katona, 1967) and teaching reasoning strategies is more effective than drill in facilitating learning, retention, and transfer of basic combinations in particular (Brownell, 1941; Brownell & Chazel, 1935; Steinberg, 1985; Swenson, 1949; Thiele, 1938; Thornton, 1978). 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. For such reasons, there is now broad agreement that instruction on the basic combinations should focus on fostering relational learning in general and reasoning strategies in particular (NMAP, 2008; NRC, 2001; Rathmell, 1978; Suydam & Weaver, 1975).

*Two of the most basic reasoning strategies.* Research indicates that two of the easiest reasoning strategies for children to learn are the add-1 rule and the near-doubles tactic (see e.g., reviews by Brownell, 1941; Cowan, 2003). Why? In his 1892 *Talk to Teachers*, William James (1958) underscored the need to build on existing knowledge to promote meaningful memorization of new knowledge (cf. Piaget, 1964): "When we wish to fix a new thing in a pupil's [mind], our conscious effort should not be so much to impress . . . it as to connect it with something already there. . . . If we attend clearly to the connection, the connected thing will certainly be likely to remain within recall" (pp. 101–102). Given the informal knowledge children bring to school, the $n + 1$ or $1 + n$ family is a developmentally appropriate (as well as logical) place to begin mental addition training. At the start of school, most pupils are so familiar with the count sequence they can fluently specify the number after a given number (e.g., "After seven comes eight"; Fuson, 1988, 1992). All they need do to achieve fluency with $n + 1$ or $1 + n$ items (i.e., learn the add-1 rule) is to connect adding with 1 to their extant number-after knowledge (Baroody, 1985, 1989b, 1992). With practice, this rule can be applied efficiently to deduce the sum of any, even a previously unpracticed, $n + 1$ or $1 + n$ item for which the counting sequence is known. The near-doubles strategy builds on prior knowledge of the easily learned add-1 rule and the doubles.[1]

*Difficulties with learning reasoning strategies and achieving fluency.* Unfortunately, many children have difficulty achieving Phase 2. In fact, a characteristic of pupils with mathematical learning difficulties is that they do not spontaneously invent reasoning strategies and independently achieve Phase 2 (Swanson & Cooney, 1985; Swanson & Rhine, 1985). As a result, achieving fluency with basic sums (Phase 3) is a serious stumbling block for many schoolchildren (e.g., Henry & Brown, 2008), and a lack of fluency is a pervasive characteristic of those who have difficulties learning mathematics (Geary, 1990, 1996, 2003; Jordan, Hanich, & Kaplan, 2003; Jordan, Hanich, & Uberti, 2003).

As achieving Phase 2 and well-structured Phase 3 knowledge seems to stem from number sense (Gersten & Chard, 1999; Jordan, 2007), inadequate informal knowledge or ineffective formal instruction (i.e., school instruction that does not build on this prior knowledge or otherwise focus on relations) may delay or prevent fluency. Children at risk for academic failure (e.g., with low early achievement or from low-income families), in particular, may have critical gaps in informal knowledge—including the developmental prerequisites of mental addition (Baroody, Eiland, & Thompson, 2009). Compounding the problem, teachers in schools with large populations of children at risk for later academic difficulties are among the most poorly trained, particularly in regards to mathematics and methods for fostering number sense (Ferguson, 1998; Lee, 2004; Lubienski, 2001, 2002; Lubienski & Shelley, 2003). Such children are particularly susceptible, then, to difficulties or delays in achieving fluency with even the most basic sums (Jordan, Huttenlocher, & Levine, 1992, 1994; Jordan, Kaplan, Olah, & Locuniak, 2006)—including $n + 1$ or $1 + n$ combinations (Baroody, Eiland, Purpura, & Reid, 2012; Purpura, Baroody, Eiland, & Reid, 2012). Effective early intervention may help many in this large and growing population achieve Phases 2 and 3 and avoid debilitating difficulties caused by not achieving fluency with basic sums (Dev, Doyle, & Valente, 2002; Fuchs et al., 2005). The focus of the present study was to evaluate the promise of computer-aided interventions with two of the most basic reasoning strategies (add-1 rule and near-doubles strategy) with first graders with a risk factor.

### First Subsidiary Aim: The Relative Value of Different Forms of Discovery Learning

Addressed, in turn, are: (a) why discovery learning may be an important means for learning reasoning strategies meaningfully and (b) what form of discovery learning may be more effective.

*Why discovery of basic reasoning strategies rather than their direct instruction?* In an extensive review of the literature, the NRC (2001) concluded that Phase 2 can be accelerated by directly teaching reasoning strategies, *if* done conceptually. Direct teaching of reasoning strategies accompanied by explanation of their rationale is often recommended by mathematics educators (Rathmell, 1978; Thornton, 1978, 1990; Thornton & Smith, 1988) and utilized in many elementary curricula, such as *Everyday Mathematics* (University of Chicago School Mathematics Project [UCSMP], 2005).

However, not all conceptually based instruction is equally effective (Baroody, Purpura, & Reid, 2012). Chi (2009) hypothesized that interactive learner activities (substantively dialoguing and considering a partner's views) are more effective in promoting learning than constructive activities (producing responses that entail ideas that go beyond provided information), which are more effective than active activities ("doing something physically"), which in turn are more effective than passive activities (e.g., listening or watching without using, exploring, or reflecting on the presented material). Direct instruction—even when it attempts to illuminate the rationale for a reasoning strategy—typically embodies passive activities. As a result, it may not actively engage many children, be comprehensible, or produce routine expertise, which leads to applying a strategy inflexibly and inappropriately. For example, Murata (2004) found that Japanese children taught a decomposition strategy with larger-addend-first combinations did not exhibit strategy transfer when smaller-addend-first items were introduced. Torbeyns, Verschaffel, and Ghesquiere (2005) found that children taught the near-doubles strategy sometimes used the strategy accurately but other times inaccurately (e.g., relating 7 + 8 to 7 + 7 – 1 or 8 + 8 + 1 instead of 7 + 7 + 1 or 8 + 8 – 1).

Discovery learning may be better suited to learning basic reasoning strategies than direct instruction because it can involve active learning and constructive activities (e.g., Swenson, 1949; Thiele, 1938; Wilburn, 1949). Alfieri et al. (2011) observed that discovery learning has increasingly supplanted traditional direct instruction in recent decades because "allowing learners to interact with materials [and] explore phenomena . . . affords them . . . opportunities to notice patterns, [relations], and learn in ways that are seemingly more robust" (p. 1). They hypothesized the more robust results may be due to a *generation effect*—the enhancement of learning and retention when learners are permitted to construct their own knowledge in some way, such as generating their own generalization (cf. Slamecka & Graf, 1978).

*What form of guided discovery learning may be more effective?* Discovery learning encompasses a wide range of methods, which may not be equally effective in all cases. At one extreme is highly guided and explicit practice (e.g., items arranged sequentially to underscore a pattern or relation, explicit scaffolding of questions or hints to guide attention to regularities, explicit feedback that explains why a response is correct or incorrect, and additional hints or scaffolding as needed). At the other extreme is unguided and implicit discovery learning (e.g., haphazardly presented items, no explicit scaffolding or feedback). Alfieri et al.'s (2011) meta-analysis of 164 studies revealed that explicit instruction was generally more effective than unguided discovery and concluded: "Unassisted discovery does not benefit learners, whereas feedback, worked examples, scaffolding, and elicited explanations do" (p. 1). Kirschner, Sweller, and Clark (2006) likewise concluded that minimally guided discovery methods were ineffective and inefficient compared to direct instructional guidance ("providing information that fully explains the concepts and procedures"). Alfieri et al. also found that enhanced or guided discovery learning was more effective than other forms of instruction (e.g., explicit instruction or unguided discovery). Hmelo-Silver, Duncan, and Chinn (2007) similarly concluded that highly scaffolded problem-based or inquiry-based instruction is effective. Although previous training studies (using an untimed mental addition task) indicate that unguided discovery learning can help typically developing kindergartners to discover the relatively salient or obvious regularities underlying the add-0 and add-1 rules and apply these rules adaptively (Baroody, 1989b, 1992), it seemed reasonable to hypothesize that guided discovery of these rules would benefit children with a risk factor more than unguided discovery.

The structured interventions to promote reasoning strategies in the present study involved minimally guided and implicit discovery learning in that children actively determined sums, reconsidered answers in the face of feedback regarding correctness only, and detected a mathematical regularity and used this regularity to construct strategies for themselves. In an effort to move beyond passive learning activities (Chi, 2009) and consistent with the generation effect (Alfieri et al., 2011) and the recommendation of the NMAP (2008), structured training involved sequentially arranging problems to highlight a relation and where that relation was applicable. For example, for the structured add-1 training, answering a number-after-$n$ question (e.g., "What number comes after 3 when we count?") was immediately followed by answering a related $n + 1$ item (e.g., "$3 + 1 = ?$"). In theory, this sequence may result in considering the $n + 1$ expression, its sum, $n$, and the number after $n$ simultaneously in working memory and thus increase the chance of inducing a localized add-1 rule. The $1 + n$ counterpart was posed next to prompt recognition of additive commutativity and the applicability of the add-1 rule to $1 + n$ items. An $n + 0$ or $0 + n$ item and an $n + m$ item (where $n$ and $m > 1$) served as nonexamples of the add-1 rule to discourage overgeneralizing the rule (i.e., to instill its appropriate use).

### Second Subsidiary Aim: Implications of a Hypothetical Learning Progression

The present research also served to evaluate the implications of an instructional/developmental model that incorporates the following two suggestions: (a) Learning trajectories and big ideas can be invaluable in guiding assessment and instructional efforts (Clements & Sarama, 2004, 2009). (b) "Future models of arithmetic [might] benefit from including retrieval structures or other mechanisms that embody numerical magnitude representations" (Siegler & Ramani, 2009, p. 556). A hypothetical trajectory and its key implications are discussed next.

*A hypothetical trajectory.* The Common Core State Standards (2011) includes as a goal for kindergarten "understand that each successive number name refers to a quantity that is exactly one larger." This goal refers to the big idea described hereafter as the *successor principle*. This principle may be the conceptual basis for (a) re-representing the counting sequence as the (positive) integer sequence ($n$, $n + 1$, $[n + 1] + 1 \ldots$) and in a linear manner (understanding that the magnitude of each successive number in the counting sequence increases by a constant amount) and (b) the add-1 rule. (The successor principle and add-1 rule are logically related but different operations. The former entails recognizing that one must be added to a number to obtain the number after it; the latter involves recognizing that the number after is the result of adding 1 to a number.) The add-1 rule may be the retrieval structure that connects children's mental addition with their representation of counting and magnitude (e.g., the successor principle and a linear representation).

*Implications.* Two implications follow from the proposed model:

1. The meaningful development of fluency with a combination family depends on first mastering family-specific developmental prerequisites. A requirement for the fluent use of both the successor principle and add-1 rule is efficient recall of number-after relations. The successor principle and add-1 rule plus automatic knowledge of the doubles are necessary for achieving fluent use of the near-doubles strategy.
2. Learning the add-1 rule may have a broad impact. A fluent add-1 rule is the bridge between computing sums concretely with objects (basic Phase 1 strategies) and doing so mentally (more advanced Phase 1 strategies and Phase 2 or 3 strategies). Specifically, this rule provides a scaffold for (a) the relatively advanced "abstract counting-on" strategy (Baroody, 1995; Bråten, 1996; Clements & Sarama, 2012), (b) inducing commutativity (Resnick, 1983), and (c) inventing more advanced reasoning strategies for larger sums such as near-doubles strategy, the skip-the-*n*-after rule for adding with 2 (e.g., $7 + 2$ is the number after the number after eight—*nine*), and the decomposition strategy (e.g., $7 + 9 = [6 + 1] + 9 = 6 + [1 + 9] = 6 + 10 = 16$). The add-1 rule and its derivative knowledge in turn are important bases for success with school mathematics. For example, early mental addition is predictive of later mathematics achievement or learning difficulties (Jordan, Kaplan, Ramineni, & Locuniak, 2009; Jordan & Levine, 2009; Mazzocco & Thompson, 2005).

## Specific Questions Addressed

The present research addressed four main questions. The first and third involve the primary aim (feasibility of interventions for promoting fluency with the most basic sums); the second, the subsidiary aim regarding the relative value of guided and unguided programs; and the fourth, the subsidiary aim regarding the implications of the model of early mental addition.

*Research Question 1:* Is computer-assisted, guided but implicit discovery or unguided discovery of the add-1 rule a feasible approach for improving first graders' fluency with $n + 1$ or $1 + n$ items? Three specific questions were addressed: (a) Can such training be more effective in promoting fluency with practiced $n + 1$ or $1 + n$ items than regular classroom instruction and practice with these combinations? (b) Might the same be true with unpracticed $n + 1$ or $1 + n$ items—a result indicative of transfer and recognition of a general add-1 rule? (c) Might intervention have a different impact on children with different levels of initial fluency?

*Research Question 2:* What type of discovery learning might be more effective in promoting first graders' fluency with $n + 1$ or $1 + n$ items? (a) Consistent with Alfieri et al. (2011) and Kirschner et al.'s (2006) reviews, is guided but implicit add-1 discovery learning more effective than unguided training in promoting fluency with practiced $n + 1$ or $1 + n$ combinations? (b) Might the same be true with unpracticed $n + 1$ or $1 + n$ items?

*Research Question 3:* Is computer-assisted, guided but implicit discovery of the connection between the doubles and the near-doubles more effective than regular classroom instruction/practice with these items in promoting first graders' fluency with practiced or unpracticed near-doubles?

*Research Question 4:* Is the evidence stemming from the training experiment consistent with the proposed learning progression of mental addition? Specifically: (a) Might achieving fluency with $n + 1$ or $1 + n$ items depend on mastering the developmental prerequisite (number-after relations), and can fluency with the near-doubles reasoning strategy be achieved with training on one prerequisite (doubles fluency) but not another ($n + 1$ or $1 + n$ fluency)? (b) Might achieving fluency with the add-1 rule have a carryover effect on the fluency of unrelated items?

## Method

### Participants

Participants were drawn from two school districts serving a midsized Midwestern U.S. community. Mental addition pretesting identified 64 first graders eligible for the study—who had not mastered more than 60% of the $n + 1$ or $1 + n$ items (range = 0% to 58%; median = 8%) or the near-doubles (range = 0% to 42%, median = 0%). All met at least one criterion for "academic risk" used by their school district: bottom 25th percentile in mathematics achievement or a personal (e.g., English as a second language, ADHD, visual impairment) or a "familial" (e.g., single-parent household, eligible for free or reduced lunch, parents without a high school diploma) factor that might place children at an increased likelihood of experiencing poor educational outcomes (e.g., grade retention, failure to complete school)—see Table 1 for details. This population was targeted because of the likelihood they would not be fluent with even the easiest basic sums and might profit from the interventions. This diverse or heterogeneous sample is reasonably representative of the extensive and growing national population of first graders with a risk factor. For the purposes of demonstrating the feasibility of the experimental interventions, the sample provides reasonable ecological and external validity (generalizability).

The three classes in School 1 ($n$ = 28) used a combination of the *Beginning School Mathematics Project* (Miller & McKinnon, 1995) and *Mathematics Their Way* (Baratta-Lorton, 1995). The nine classes in Schools 2 to 4 ($n$ = 36) used *Everyday Mathematics* (UCSMP, 2005). All programs include activities for both group and individual work with manipulatives and materials common to early childhood classrooms, but none included instructional software. Although teachers provided computer time to play math games, given the scarcity of structured discovery software for young children, it is likely these programs focused on drill of number skills. All schools were committed to achieving the state's Grade 1 objectives (http://www.isbe.net/ils/math/standards.htm), which included addition and subtraction with whole numbers (e.g., solving one- and two-step problems and performing computational procedures). Only the *Everyday Math* program made an effort to teach the add-1 rule by relating $n + 1$ to the number after $n$ (see Table 2). However, this involved only some explicit or direct instruction by teachers and practicing number-after relations and $n + 1$ or $1 + n$ items separately or implicitly (e.g., adding 1 using a number line).

Project-hired personnel—one male and four female academic professionals and two male and three female research assistants—were highly qualified to teach and test young children and to observe the learning process (provide qualitative data for formative assessment of the experimental programs). All had previous teaching experience, training in education, or in most cases both, and eight had worked with pre-K children the previous year of the project. Project staff received extensive training—six 3-hour training sessions on testing and training procedures.

### Measures

The achievement/diagnostic test and mental addition test used are described in turn.

*Achievement and diagnostic testing.* The Test of Early Mathematics Ability–Third Edition (TEMA-3; Ginsburg & Baroody, 2003) is a manually and individually administered, nationally standardized test of math achievement for 3- to 8-year-olds. The test measures informal and formal concepts and skills in the following domains: numbering, number comparison, numeral literacy, combination fluency, and calculation. Cronbach's alphas for the form used (Form A) overall is .94 and the three age intervals

*Table 1*
**Sample Characteristics and Mental Addition Test Results by Condition**

| Training Condition | Structured Add-1 ($n$ = 21) | Unstructured Add-1 ($n$ = 22) | Structured Near-Doubles ($n$ = 21) |
|---|---|---|---|
| Age range/median age | 6.1-7.4/6.5 | 6.1-7.8/6.6 | 6.0-7.3/6.4 |
| Males | 10 | 5 | 8 |
| TEMA-3[a] | | | |
| Pretest range (*SD*) | 62-113 (11) | 62-99 (10.6) | 65-102 (12.9) |
| Pretest mean (median) | 91 (89) | 85 (86) | 83.8 (89) |
| Posttest range (*SD*) | 71-117 (10.5) | 72-100 (7.8) | 72-107 (11.5) |
| Posttest mean (median) | 91.7 (92) | 88.2 (88) | 88.4 (87) |
| Risk factors | | | |
| Math achievement < 25th percentile | 11 | 15 | 11 |
| Free or reduced lunch | 12 | 16 | 17 |
| Single-parent family | 7 | 11 | 8 |
| Parent under 18 years old | 0 | 4 | 2 |
| Parents without a high school diploma | 0 | 1 | 3 |
| Black/Hispanic/multiracial/English as a second language | 14 | 17 | 18 |
| Birth complications | 0 | 0 | 4 |
| Low birth weight | 1 | 1 | 2 |
| Fetal alcohol/drug syndrome | 0 | 0 | 1 |
| Visual impairment | 0 | 0 | 1 |
| Speech services | 2 | 4 | 2 |
| ADHD | 3 | 2 | 1 |
| Aggressive/acting out | 1 | 3 | 3 |
| Passive/withdrawn | 0 | 3 | 2 |
| Mental-addition fluency: Practiced items | | | |
| Pretest mean (*SD*) for $n$ + 1/1 + $n$ items | .20 (.24) | .15 (.21) | .15 (.18) |
| Posttest 1 mean (*SD*) for $n$ + 1/1 + $n$ items | .67 (.34) | .65 (.27) | .33 (.36) |
| Posttest 2 mean (*SD*) for $n$ + 1/1 + $n$ items | .68 (.39) | .73 (.29) | .48 (.39) |
| Pretest mean (*SD*) for near-doubles | .04 (.11) | .02 (.05) | 0 (0) |
| Posttest 1 mean (*SD*) for near-doubles | .13 (.24) | .05 (.10) | .10 (.15) |
| Posttest 2 mean (*SD*) for near-doubles | .14 (.24) | .07 (.10) | .14 (.16) |
| Mental-addition fluency: Unpracticed items | | | |
| Pretest mean (*SD*) for $n$ + 1/1 + $n$ items | .18 (.30) | .19 (.26) | .19 (.27) |
| Posttest 1 mean (*SD*) for $n$ + 1/1 + $n$ items | .60 (.41) | .52 (.40) | .42 (.43) |
| Posttest 2 mean (*SD*) for $n$ + 1/1 + $n$ items | .66 (.41) | .66 (.31) | .46 (.43) |
| Pretest mean (*SD*) for near-doubles | .04 (.09) | .02 (.07) | 0 (0) |
| Posttest 1 mean (*SD*) for near-doubles | .06 (.13) | .07 (.16) | .02 (.08) |
| Posttest 2 mean (*SD*) for near-doubles | .10 (.22) | .10 (.18) | .08 (.16) |

*Note.* The schools provided this information and, where parent permission was granted, project staff double-checked a child's school files.
[a]Children in the three conditions did not differ significantly on the Test of Early Mathematical Ability–Third Edition (TEMA-3) pretest, $F(2, 61) = 2.33$, $p = .106$ or TEMA-3 posttest, $F(2, 61) = .78$, $p = .461$.

represented by the sample (5- to 7-year-olds) are .94 or .95. 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-prediction validity, correlations between

*Table 2*

**Summary of the Mental Addition Training and Practice Provided in Each of the Schools**

| Curriculum | General | $n + 0$ or $0 + n$ Items | $n + 1$ or $1 + n$ Items | Doubles |
|---|---|---|---|---|
| *Beginning School Mathematics Project* (Miller & McKinnon, 1995) (School 1) m | Using sets no larger than 7, students are expected to:<br>- Identify the set that comes just after, just before, or between two given sets.<br>- Demonstrate understanding of the "one more than" and "one fewer than" relations.<br>- Join two sets and discuss the operation informally. | No explicit analogies for +0 items are given. | No explicit analogies for +1 items are given. | No explicit strategy |
| *Mathematics Their Way* (Baratta-Lorton, 1995) (School 1) | Addition combinations up to 18 practiced using a variety of games. | No explicit analogies for n + 0 items are given. | No explicit analogies for n + 1 items are given. | Doubles are practiced but no explicit strategy is presented so as assist in their calculation. |
| *Everyday Mathematics* (University of Chicago School Mathematics Project, 2005) (School 2) | Spends weeks on memorizing the "easy facts" (e.g., Unit 2 Lesson 3, Unit 4 Lessons 11 and 12, Unit 5 Lessons 10 and 11; Unit 6, Lesson 1): $n + 0$, $n + 1$, and doubles and items that sum to 10. The $0 + 0$ to $9 + 9$ items are practiced often (e.g., Unit 3 Lesson 14 and Unit 4 Lesson 11 use dominoes); the notion of a turn-around fact (the same numbers are being added so they have the same sum) is introduced to lessen the number of addition combinations to be learned (e.g., Unit 5 Lesson 11). Worksheets with $n + 0$ in row 1, $n + 1$ in row 2, and doubles along a diagonal—each shaded in a different color—assigned repeatedly.<br>Information cited in this table appears in Teacher's Manual but not student worksheets. | $N + 0$ is related to the sum is always the same as the number you start with. | $N + 1$ is related to the sum is always the number that comes after the number you start with.<br>Unit 1 Lesson 1 introduces 1 more by practicing number-after and adding 1 separately or implicitly (e.g., using a number line to add 1).<br>In Unit 3, Lesson 8, a teacher explains: "Counting up by 1s is like adding 1 to each number to get the next number." | Doubles are highlighted as the numbers along the diagonal and are always even. Skip counting, such as counting by twos (2, 4, 6…) and fives (5, 10, is practiced (e.g., Unit 3 Lesson 9) |

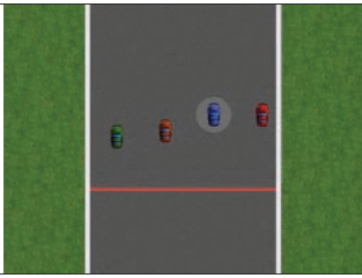
*Note*. None of the curricula used in the schools focused on the near-doubles.

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. The following TEMA-3 items served to assess developmental prerequisites for mental addition and were administered to all participants for diagnostic purposes: No. 11 (produce fingers to 5), No. 13 (number after to 9 + extra trials of $n = 2$, 4 and 6), No. 16 (concrete modeling of addition), No. 17 (part-part-whole), No. 19 and No. 20 (number comparisons to 5 and 10, respectively), No. 23 (counting sets to 10), No. 26 (mental addition: sums 5 to 9), No. 31 (verbally counting by ones to 42), and No. 32 (counting-on).

*Mental addition test.* The testing was done in the context of computer games developed for the project. Figure 1 illustrates the Race Car Game and summarizes details on the computer-assisted testing procedure. If a child did not respond within 5 seconds, the tester prompted: "What do you think would be a good guess?" The problem was also placed in context of a word problem (e.g., "Three cookies and one more cookie is about how many cookies altogether?"). The test of mental addition fluency included six categories of items: (a) practiced in the add-1 conditions (0 + 4, 0 + 7, 3 + 0, 8 + 0, 1 + 3, 1 + 4, 1 + 7, 1 + 8, 3 + 1, 4 + 1, 7 + 1, 8 + 1), (b) unpracticed (transfer) items for the add-1 condition (0 + 6, 9 + 0, 1 + 6, 1 + 9, 6 + 1, 9 + 1), (c) practiced in all conditions (3 + 5, 5 + 7), (d) transfer items for all conditions (5 + 3, 7 + 5), (e) practiced in the near-doubles condition (2 + 2, 3 + 3, 4 + 4, 5 + 5, 6 + 6, 8 + 8, 2 + 3, 3 + 4, 5 + 6, 6 + 7, 3 + 2, 4 + 3, 6 + 5, 7 + 6), and (f) transfer items for the near-doubles condition (4 + 5, 8 + 9, 5 + 4, 9 + 8).

Each testing session consisted of a test set of 10 items, computer reward game, second test set of 10 items, and final reward game. The mental addition items were presented in a partially random order. Specifically, the order was haphazard, except that two items with the same addends or sum were not presented one after another, commuted items were not presented in the same session, and the types of items were evenly distributed across the four sets.

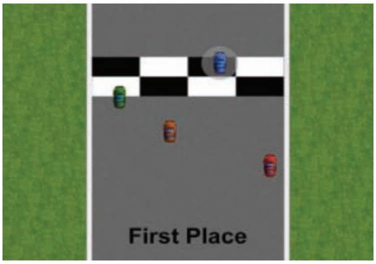
Item fluency was defined as an efficient response (i.e., the correct sum in less than 3 seconds) without evidence of counting or a false positive due to a response bias (that some mental addition novices are prone to use).[2] A response bias was defined as applying a strategy to at least half of the applicable and nonapplicable items alike during a testing session. For example, nonselective use of the add-1 rule was inferred for session 2 if a child stated the number after the larger addend for 7 or more of the 14 $n + 0$ or $0 + n$, doubles, near-doubles, and filler items and 3 or more of the 6 $n + 1$ or $1 + n$ trials. This straightforward scoring procedure yielded 100% interrater agreement on fluent versus nonfluent scores on all $n + 1$ or $1 + n$ items for 7 randomly chosen participants in each of the three conditions for each of the three testing sessions. (For a detailed discussion of the scoring



a. The computer highlighted a child's choice of a racecar. The game provided a pretext for tightly holding onto a steering wheel to eliminate or at least suppress finger counting. The child was encouraged to respond quickly and, if not sure of an answer, to quickly make a good guess After the testing instructions were given, the tester pressed the ENTER key to start the race. The racecars sped across the starting line, and—after a moment—the screen faded to black.


b. After a blank block appeared to provide a focus for the child's attention, the tester keyed in ENTER on the number pad to initiate a trial and to start the computer's internal stopwatch. Children were not provided objects or encouraged to use objects during the mental-addition testing. As soon as the child responded, the tester (who had his/her finger posed on the ENTER key) depressed the key to stop the internal stopwatch.


c. The child's RT was displayed in the upper left-hand corner and filed electronically. Using the number pad, the tester then entered the child's response and hit ENTER to record the answer electronically. The entered response appeared on the screen in the upper right-hand corner. The tester also recorded the RT and answer manually as a backup and any relevant information (e.g., a response before the trial was displayed, counted objects, finger counting, verbal counting, reasoning such the number after 7 is 8 or 4 and 4 is 8 and 1 more 9).




d. After the testing, the race picked up where it left off and continued for about 15 s. Where the child's car finished was determined by the quality (accuracy and speed) of his/her responses during the testing session. For the purposes of this feedback (but not the scoring of fluency), items were scored on accuracy—correct, close (within 25% of the correct sum if > larger addend), reasonable (within 150% of the correct sum and larger than the addends), or unreasonable—and speed (fast or slow). The average score determined the outcome of the game (e.g., the place the child's car finished in the race).

*Figure 1.* **Example of mental addition testing game (Race Car Game).**

procedures used in this study and the resulting data on response biases, see Baroody, Purpura, Reid, & Eiland, 2011.)

### Interventions

Bruner (1961) cautioned that discovery learning requires base knowledge or developmental readiness. Consistent with the number sense view (Baroody, 1985; Baroody, Bajwa, et al., 2009; Brownell, 1935; Gersten & Chard, 1999; Jordan, 2007) and recommendations stemming from research reviews (e.g., Clements & Sarama, 2012; NMAP, 2008; NRC, 2001; see also Hasselbring, Goin, & Bransford, 1988), the present study involved an initial training stage (Stage I) that included Phase 1 learning experiences (using counting strategies to solve word problems) to ensure adequate base knowledge for the three experimental programs (Stage II): recognizing relations and constructing and automating reasoning strategies (Phases 2 and 3 learning experiences). Stage I training and Stage II interventions supplemented regular classroom mathematics instruction and are described in turn.

*Stage I.* Stage I training focused on the prerequisites for mental addition needed by all groups. The goals and sequence of this training are summarized in Table 3. The sequence starting point and subsequent training was tailored to the individual developmental needs of a child.

Training occurred in the context of manual games, which typically involved several goals. For example, Animal Spots involved throwing a die or dice to determine how many pegs (spots) a player could take on his or her turn to fill holes in a wooden cutout of a leopard or giraffe. (The player who filled his or her animal with "spots" first was the winner.) A dot die was used to practice verbal number recognition (subitizing) or enumeration; a dot and numeral die was used to connect numerals to concrete collections, the numeral die served to practice numeral recognition, two dot dice were used to introduce addition and practice a counting-all strategy, two dot-plus-numeral dice were used to introduce symbolic addition, and two numeral-only dice were used to practice symbolic addition. This game also entailed practicing the cardinality rule (the last number used in counting a collection represents the total) and verbal production (counting out a specified number of pegs). Racecar, which entailed rolling one die or two dice or drawing a card, had the same goals as those just described. Its Number-After version also served to practice the number-after skill. For example, if a child turned up the after 3? card from a deck of cards (and—if necessary—read by the trainer) and answered the number-after question correctly, he or she moved his or her race car the number of spaces equal to the answer. A primary source of the games was Baroody (1987, 1989a, 1998) and Wynroth (1975/1986). A detailed description of the games and their sequence is available from the first author.

*Stage II.* Although little software exists to help young children discover relations or invent reasoning strategies, such software might be particularly valuable for early intervention with at-risk children (Rasanen, Salminen, Wilson, Aunio, & Dehaene, 2009). Well-designed computer programs and

*Table 3*
**Goals of Stage I of the Intervention and Their Difficulty Levels**

| Goals | Difficulty Level 1 | Difficulty Level 2 | Difficulty Level 3 |
|---|---|---|---|
| **Basic number** | | | |
| Nonverbal matching: using a visible collection to create a matching collection | 1 to 3 | 1 to 6 | |
| Nonverbal set production: creating (producing) a matching collection for a previously seen but then hidden collection | 1 to 3 | 1 to 4 | 1 to 6 |
| Verbal number recognition (VNR): immediately, accurately, and reliably recognizing the number of items in (small) collections | 1 to 3 | 1 to 4 | 1 to 6 |
| Using VNR to put out a requested *n* (basic give to me-*n* task) | 1 to 3 | | |
| **Verbal counting** | | | |
| Verbal counting: generating the count words in the correct sequence up to 21 | 1 to 10 | 1 to 12 | 1 to 20 |
| Number after: automatically citing the *number after a term* without a running start (without counting from 1) | 1 to 5 | 1 to 10 | 1 to 20 |
| **Object counting** | | | |
| Enumeration: 1-to-1 counting | 1 to 5 | 1 to 10 | 1 to 20 |
| Cardinality rule: recognizing that the last number word used in the enumeration process represents the total | | | |
| Counting out a requested *n* (advanced give-me-*n* task) | 1 to 5 | 1 to 10 | 1 to 20 |
| **Numerical relations** | | | |
| Intuitive (perception-based) more/“same” comparisons | | | |
| Same number based on VNR | 1 to 3 | 1 to 4 | 1 to 6 |
| Gross mental “more”/ “fewer”: recognizing which of two widely separated number words represents the larger or smaller collection | 1 to 5 | 1 to 10 | |
| Counting-based “same”/“more”/”fewer” comparisons: using verbal or object counting to determine numerical relations of two collections or numbers | 1 to 5 | 6 to 10 | |
| Close mental comparisons of “more”/“fewer”: recognizing which of two consecutive numbers represents the larger or smaller collection | 1 to 5 | 6 to 10 | |
| **Written numbers** | | | |
| Numeral recognition: accurately and reliably identifying or reading numerals (written numbers) | 1 to 5 | 1 to 10 | 1 to 20 |
| Numeral writing | 1 to 5 | 1 to 10 | 1 to 20 |

*(continued)*

*Table 3* **(continued)**

| Goals | Difficulty Level 1 | 2 | 3 |
|---|---|---|---|
| **Arithmetic** | | | |
| Nonverbal ± concepts: mentally represent a small, previously seen collection and accurately determine the result of adding an item or a few items to it or taking away an item or a few items | to 3 | to 6 | |
| Concretely counting to add: counting two collections (e.g., the dots on two dice) to determine the total (sum) | to 5 | to 10 | |
| Estimation of sums: basic estimation skill includes recognizing that addition makes an initial collection larger (e.g., indicating that 3 + 2 is 5) | to 5 | to 10 | |
| Concrete counting all (concretely modeling addition): using objects or fingers to concretely represent addition situations (word problems or arithmetic expressions such as 3 + 2) and to determine the sum | to 5 | to 10 | |

properly chosen computer games can provide effective instruction and practice even for young children (Clements & Sarama, 2012; NRC, 2009; Sarama & Clements, 2009). A program can provide the scaffolding for guided discovery learning that most teachers cannot. Specifically, it can underscore relations, such as the connection between number-after relations and adding 1 (see Figure 2 for details), a connection that may not be known, explained clearly, or emphasized by most teachers. The game context provided motivation to learn for children accustomed to being entertained by television and online/mobile/computer games.

The three training conditions (structured add-1, unstructured add-1, and structured near-doubles) had parallel structures and identical dosage and involved the same reward games. All were done in the context of two computer games: Castle Game and Train Game (see Figures 2 to 4). All participants were initially encouraged to determine sums independently by using mental arithmetic. All practiced the correct answer to their assigned practiced items and the filler items (3 + 5 and 5 + 7) the same number of times. Both the computer and the trainer provided feedback regarding correctness. If a child responded incorrectly to an item, he or she was instructed to try again and, if incorrect again, encouraged or helped to use physical manipulatives to determine the correct. In each session, the child completed a subset of 10 items, took a break, usually in the form of a brief manual game, and completed another subset of 10 items. In the few cases where children were repeatedly not able to provide reasonable estimates, mental addition

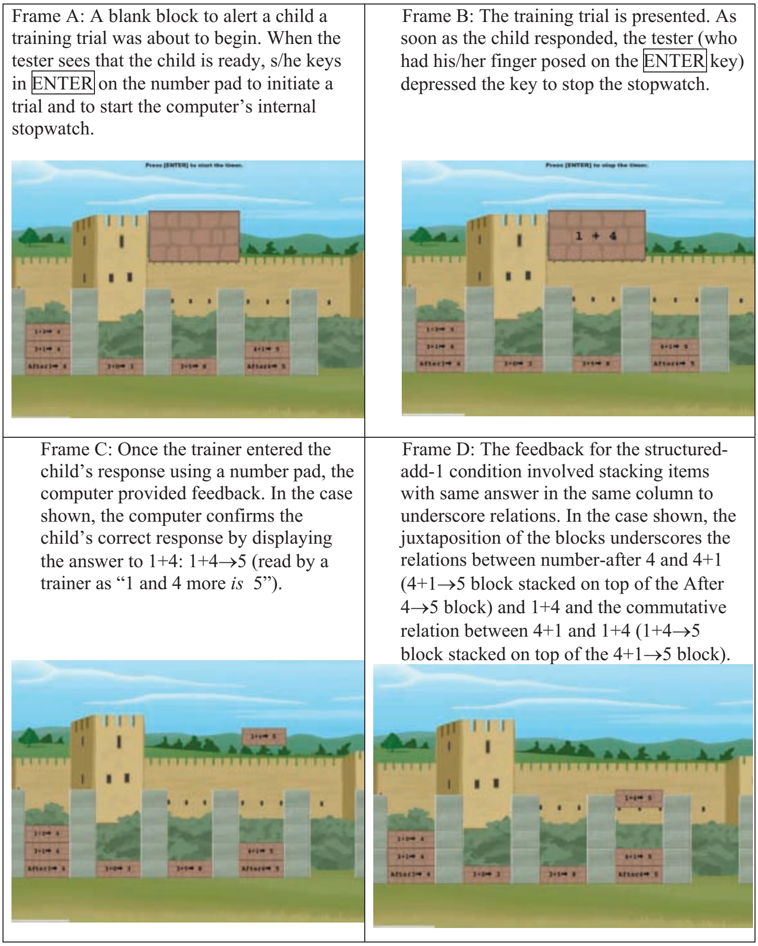



*Figure 2.* **Structured add-1 training: Castle Wall Game.**

practice was provided with a nontested item. A computer reward game was played at the end of a session.

The structured add-1 intervention provided an opportunity for minimally guided and implicit discovery in two ways. One feature was the *ordering* of the related number-after and add-1 items. For example, Set 1 consisted of the following trials in the order delineated: after 3, 3 + 1, 1 + 3, 3 + 0, 3 + 5; after 4, 4 + 1, 1 + 4, 0 + 4, after 6; Set 2: after 7, 7 + 1, 1 + 7, 0 + 7, 5 + 7; after 8, 8 + 1, 1 + 8, 8 + 0, after 9. A second feature


*Figure 3.* **Structured add-1 training: Train Game.**

*Note.* Display shows feedback at the end of a subsession. Note that all items with the same answer are on the same track. After all trials were completed, each train rode off moving left to right.

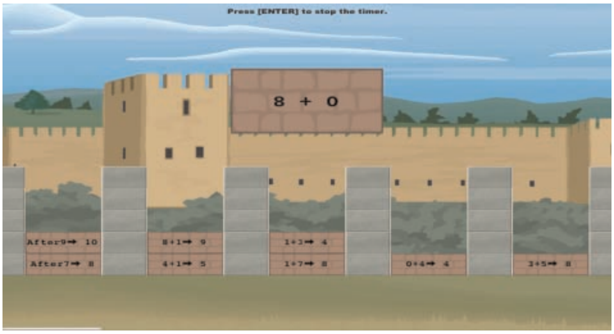



*Figure 4.* **Unstructured add-1 training: Castle Wall.**

*Note.* In the unstructured add-1 condition, items were presented in a semi-haphazard order and the feedback involved stacking items from left to right in the order in which they are presented (i.e., one item was added to each column from left to right, and then this process was repeated). As shown in the figure, for example, the After 7→8 trial was presented first, the 4 + 1→5 was presented second, and the 1 + 7→8 trial was presented third.

was feedback involving *juxtaposing related equations* (e.g., see Frame D of Figure 2). Neither the trainers nor the program explicitly pointed out the relation between adding one and number after (e.g., feedback focused on correctness only).

The training for unstructured add-1 condition was identical to that of the structured add-1 condition but essentially involved unguided discovery. That is, other than providing extra practice of number-after relations and add-1 items and feedback about the correctness of a pupil's answer, it provided no scaffolding. Specifically, items were practiced in an order designed to be "relationally haphazard"—an unrelated item always followed an item. For instance, Set 1 entailed the following trials: after 6, 1 + 4, 3 + 0, 7 + 1, 5 + 7; 1 + 8, after 4, 3 + 1, 0 + 7, after 8; Set 2: after 7, 4 + 1, 1 + 7, 0 + 4, 3 + 5; after 9, 8 + 1, 1 + 3, 8 + 0, after 3. Feedback regarding the correctness of add-1 items was not juxtaposed with that for relevant number-after relations.

The structured near-doubles condition entailed presenting a double (2 + 2, 3 + 3, 5 + 5, or 6 + 6), a related three-addend item (e.g., 2 + 1 + 2), a related near-double (3 + 2, 4 + 3, 6 + 5, or 7 + 6), the commuted partner (e.g., 2 + 3), and another double (4 + 4 or 8 + 8) or filler item (3 + 5 or 5 + 7).

### Design and Procedures

The research plan/procedures and then the rationale for the design are discussed in turn.

*Plan/procedures.* A training experiment with multiple baselines served to evaluate the feasibility of the programs. A manual TEMA-3 pretest in September gauged mathematics achievement and identified specific strengths/weaknesses in informal/formal concepts and skills. The manual Stage I intervention from October to January focused on remedying knowledge gaps identified by the TEMA-3, particularly the developmental prerequisites for mental addition, such as number-after relations. This training was provided to all participants in pairs.

The computer-based mental-addition pretest was then individually administered in January. Participants were then randomly assigned to a condition for one-on-one, computer-assisted Stage II intervention from February to mid-April. Stages I and II were *each* administered for 30 minutes twice a week for 10 weeks in a hallway outside the children's classroom. Project personnel were involved in each of the three Stage II training conditions, which were conducted simultaneously. Mental addition fluency was assessed immediately after the training at the end of April and 2 weeks after the immediate posttest in May. The latter served to assess retention. The TEMA-3 was readministered between the mental addition posttests in May. Project personnel implemented all testing and training procedures. Positive assent was obtained for all testing and training sessions. Pull outs occurred in nonliteracy time blocks, including mathematics instruction and play time. Project computer stations with an Apple G4 desktop were used for the mental addition pretest/posttests and Stage II training.

*Rationale.* As the structured near-doubles condition did not involve practice with number-after relations or the $n + 1$ or $1 + n$ items, it served as active control group for the add-1 conditions to evaluate whether minimally guided and implicit discovery and/or essentially unguided discovery of the add-1 rule was more effective than classroom instruction and practice with this combination family. In turn, the add-1 groups, which did not practice the doubles or near-doubles, served as an active control for the structured near-doubles intervention to evaluate whether the latter was more effective than regular classroom training.

The unstructured add-1 condition, which involved supplemental practice of $n + 1$ and $1 + n$ items, served as a comparison group for the structured add-1 condition, which involved supplemental practice of an $n + 1$ or $1 + n$ item *plus* modest scaffolding (juxtaposed practice and feedback of a number-after relation and associated add-1 item). The comparison of these two conditions served to evaluate whether supplemental practice alone (essentially unguided discovery) was sufficient to promote fluency with $n + 1$ or $1 + n$ items or whether some scaffolding (minimally guided discovery) was necessary to ensure supplemental practice was effective.

Various threats to internal validity are controlled by the design. Specifically, significant posttest differences cannot be attributed to history (e.g., classroom instruction or practice), maturation (e.g., natural growth of working memory capacity), regression to mean, or selection, because random assignment theoretically ensures all groups are comparable on these confounding variables. A testing effect (learning merely due to repeated testing) can be discounted because all groups receive the identical tests the same number of times. As all groups received their training via computer and reward games, the design also controls for a novelty effect. This also eliminates as a plausible alternative explanation that any effect is merely due to computer-based delivery and not the structured training or practice. Any contamination or diffusions effect, which would facilitate the learning of the $n + 0$ or $0 + n$, $n + 1$, or $1 + n$ items (e.g., sensitizing a child to mathematical regularities in general or the number-after-$n$ rule for adding 1 in particular) adds to measurement error and makes it more difficult to obtain significant results (i.e., effectively stacks the deck against corroborating a feasibility hypothesis).

*Fidelity.* Developmentally appropriate and needed Stage I instruction was assured by a checklist (Table 3) of TEMA-identified strengths and weaknesses for each child and a manual of specific games or game variations for each goal. Project staff updated the checklist as a child progressed. Fidelity of the Stage I and II training was ensured by (a) devoting 10 hours of the staff training to a learning trajectory of mental addition and its developmental prerequisites (see Baroody, Eiland, et al., 2009), the rationale and rules for the Stage I manual games, the rationale for the Stage II programs, procedures for implementing the computer-assisted Stage II training and providing feedback, and behavioral management techniques; (b) oversight (on average a session every other week for each staff member) by the Project Director or Head Project Teacher; (c) brief (10 to 30 minute) staff meetings during the training to review procedures and address training issues as needed; and (d) a lesson log sheet for keeping track of which lessons each participant completed. Stage II fidelity was further ensured by keeping a copy of the *Trainer Guidelines* at each computer station and, most importantly, the use of the computer programs. The programs ensured that each child received (a) the assigned intervention (a child's log-in automatically connected to his or her treatment), (b) the items in the order specified by an intervention, and (c) immediate feedback on correctness. In the add-1 and doubles conditions, the programs also ensured that the child saw feedback that juxtaposed elements of a relation (e.g., 3 + 1 = 4 and after 3 is 4). In regard to implementation fidelity, all participants completed 100% of the lessons in their assigned Stage II training before posttesting.

### Analytic Procedures

Analyses used a participant's mean proportion of a combination family/category scored as fluent. A series of 2 (condition) × 3 (time) repeated measures ANOVAs were used to compare children's improvement on targeted practiced and unpracticed combinations. Planned contrasts were used to compare children's performance at pretest, immediate posttest, and delayed posttest. Effects of treatment were tested using one-tailed significance values given the directional nature of the contrasts (structured add-1group > unstructured add-1 group, combined add-1 groups > control [near-doubles] group, and near-doubles group > control [add-1] group; Knotternerus & Bouter, 2001; Pillemer, 1991). As random assignment was not done within class, the effect of school and class instruction/teacher was checked. When school and class were included as fixed-effect variables, all primary contrasts failed to show significant variance at the school or class level. As there were no significant school- or class-level effects on children's performance, only the results of the repeated measures ANOVA are presented. All assumptions for the ANOVAs were met (see e.g., Maxwell & Delaney, 2004), except for the analyses involving filler items. However, ANOVAs are robust against violations when the sample sizes of groups are comparable. Although fluency and TEMA-3 pretest results did not have a normal distribution, they did have distributions that were expected. A correction (Benjamini & Hochberg, 1995) was applied to correct for Type I error due to multiple comparisons. A total of four comparisons were included in the correction (i.e., comparison of groups at immediate posttest and delayed posttest on practiced items and again on unpracticed items).

We also examined effect size magnitude (Cohen's *d*) for all specific contrasts of interest due to the limited power of the study and the importance of evaluating effect sizes (Wilkinson & APA Task Force on Statistical Inference, 1999). A difference-in-difference adjustment was used to correct effect sizes for non–statistically significant pretest differences between groups. According to Cohen (1992), a *d* of .20 is a small effect, .50 is a medium effect, and .80 is a large effect. An effect size that is equivalent to a Hedge's *g* of at least .25 meets the criterion for effective practice set by the federal What Works Clearinghouse (Institute of Education Sciences [IES], 2011). All reported *d*s of at least .25 met this criterion.

## Results

The mean proportion (and standard deviations) of practiced and unpracticed combinations scored as fluent by condition, combination type (family), and test is detailed in Table 1. Two participants in the guided add-1 condition had relatively high TEMA-3 pretest scores of 110 and 113. Analyses conducted without these outliers did not yield appreciably different results. Therefore, the analyses for each question reported in the following include all participants.

### Question 1: Feasibility of Add-1 Training

The first set of analyses was directed at addressing a primary aim—answering Questions 1a and 1b: (Might children who received either add-1 training perform significantly better on practiced and unpracticed $n + 1$ or $1 + n$ items at the immediate posttest and at the delayed posttest than the children who received near-doubles training?) and Question 1c (Might children with relatively low or high pretest fluency benefit more from the add-1 training?).

*Fluency on practiced and unpracticed* n + *1 or 1* + n *items.* Mauchley's test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 6.93$, $p = .031$, for the practiced items but was met for the unpracticed combinations. Therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon = .90$) for the former but not the latter. Main effects of time indicated that children made significant improvements in fluency on the practiced items, $F(1.80, 110) = 109.34$, $p < .001$, and the unpracticed combinations, $F(2, 124) = 44.94$, $p < .001$. Interaction contrasts indicated that the participants in the add-1 interventions had significantly greater gains in fluency than those in the active control group with practiced $n + 1$ or $1 + n$ items, $F(3.61, 11.10) = 5.20$, $p < .001$, and marginally significantly greater gains in fluency than those in this control group with unpracticed $n + 1$ or $1 + n$ combinations, $F(4, 122) = 1.64$, $p = .085$. Pretest results were not significantly different for practiced items, $F(.90, 55.02) = .21$, $p = .646$ (two-tailed), or unpracticed combinations, $F(1, 61) = .00$, $p = .952$ (two-tailed). For practiced $n + 1$ or $1 + n$ items, the intervention groups performed significantly better than the active-control group at the immediate posttest, $F(.90, 55.02) = 14.96$, $p < .001$, $d = .90$, and at the delayed posttest, $F(.90, 55.02) = 6.27$, $p = .008$, $d = .56$. For unpracticed items, although add-1 participants did not significantly outperform those in the comparison condition on the immediate posttest, $F(1, 61) = 1.67$, $p = .101$, the effect size ($d = .34$) exceeded the .25 criterion for effective practice set by the federal What Works Clearinghouse (IES, 2011). Moreover, the results on the delayed posttest were significant, $F(1, 61) = 3.53$, $p = .033$, $d = .50$. All significant results remained so after applying Benjamini-Hochberg corrections (Benjamini & Hochberg, 1995).

As Schools 2 to 4 used *Everyday Mathematics* (UCSMP, 2005), which if implemented with any fidelity involved direct instruction on the add-1 rule, the add-1 participants in these schools were compared to their peers in the comparison group to see if the experimental add-1 training produced a substantial benefit above and beyond classroom training. An ANCOVA with pretest score as the covariate revealed that after using the Benjamini-Hochberg correction, the add-1 pupils were not significantly higher on the practiced $n + 1$ or $1 + n$ items for the immediate posttest, $F(1, 33) = 4.36$, $p = .023$, $d = .62$, or the delayed posttest, $F(1, 33) = 1.40$, $p = .123$, $d = .38$. However, both effect sizes exceeded the .25 criterion for effective practice (IES, 2011). The same was true for the unpracticed $n + 1$ or $1 + n$ items at the immediate posttest, $F(1, 33) = .74$, $p = .199$, $d = .26$, and the delayed posttest, $F(1, 33) = 1.85$, $p = .092$, $d = .42$.

*Pretest Fluency × Condition.* Given the significant effects of the add-1 training, four separate stepwise regression analyses using pretest $n + 1$ or $1 + n$ scores, intervention condition (structured + unstructured add-1 vs. near-doubles), and a Pretest × Condition interaction term were conducted to determine if this training differentially benefitted children with different initial $n + 1$ or $1 + n$ fluency. Although the pretest fluency ($\beta = .77$, $p < .001$) and the intervention condition ($\beta = .41$, $p < .001$) significantly predicted immediate posttest scores for practiced $n + 1$ or $1 + n$ combinations, no significant differences were evident for the Pretest × Condition interaction term ($\beta = -.296$, $p = .120$). The pretest scores ($\beta = .84$, $p < .001$) and intervention condition ($\beta = .26$, $p = .014$) significantly predicted delayed posttest scores for practiced $n + 1$ or $1 + n$ items, and the Pretest × Condition interaction (see Figure 5) was significant ($\beta = -.40$, $p = .049$). After the application of the Benjamini-Hochberg correction, this interaction was found to be marginally significant. The pretest scores ($\beta = .86$, $p < .001$) significantly predicted immediate posttest scores for unpracticed $n + 1$ or $1 + n$ combinations. The intervention condition ($\beta = .17$, $p = .099$) and $n + 1$ or $1 + n$ Pretest × Condition interaction term ($\beta = -.34$, $p = .059$) were marginally significant predictors. Although the pretest scores ($\beta = .70$, $p = .001$) and condition ($\beta = .23$, $p = .040$) were significant predictors of delayed posttest scores for unpracticed $n + 1$ or $1 + n$ items, the Pretest × Condition interaction (see Figure 6) was not significant ($\beta = -.26$, $p = .173$).

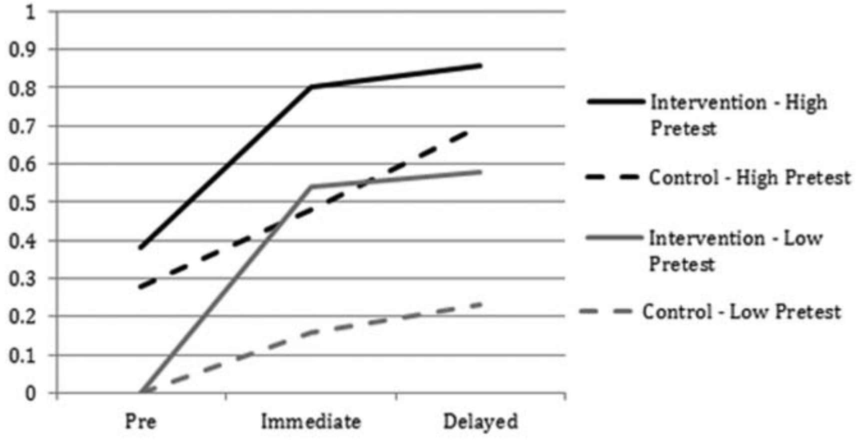



*Figure 5.* **Condition × Pretest Fluency interactions for practiced *n* + 1/1 + *n* items.**

### Question 2: Comparison of Guided and Unguided Discovery

The second set of analyses served to address the subsidiary aim of whether children who received guided add-1 training performed significantly better on practiced and unpracticed $n + 1$ or $1 + n$ items at the immediate posttest and at the delayed posttest than the children who received unguided add-1 training. Sphericity assumptions were met in both cases. Main effects of time indicated that participants made significant improvements in their fluency with practiced $n + 1$ or $1 + n$ items, $F(2, 82) = 102.98$, $p < .001$, and unpracticed items, $F(2, 82) = 46.02$, $p < .001$. Interaction contrasts indicated no differential gains between the guided and unguided groups in their fluency with practiced or unpracticed $n + 1$ or $1 + n$ items, $F(2, 82) = .64$, $p = .264$ and $F(2, 82) = .42$, $p = .329$, respectively, and thus, specific contrasts at each time-point are not reported.

### Question 3: Feasibility of Near-Doubles Training

The third set of analyses served to address a primary aim—whether the guided near-doubles training was feasible. Mauchley's test indicated that the assumptions of sphericity had been violated for the practiced and unpracticed near-doubles, $\chi^2(2) = 24.11$, $p < .001$ and $\chi^2(2) = 14.88$, $p = .001$, respectively. Therefore, degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\epsilon = .75$ and $.82$, respectively). Main effects of time indicated that children made significant improvements in their fluency with practiced near-doubles, $F(1.50, 91.67) = 20.31$, $p <$

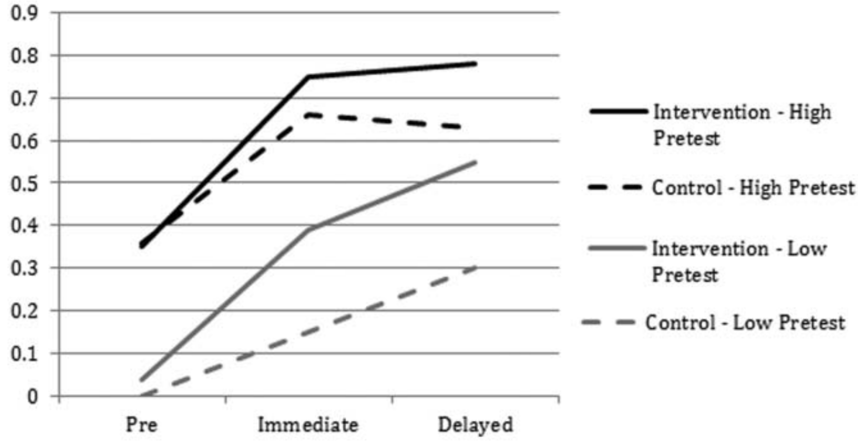



*Figure 6.* **Condition × Pretest Fluency interactions for unpracticed *n* + 1/1 + *n* items.**

.001, and unpracticed near-doubles, $F(1.64, 100.03) = 8.50$, $p = .001$. Interaction contrasts indicated marginally significantly greater gains for the intervention group than for the active-control group for the practiced near-doubles, $F(3.01, 91.67) = 1.66$, $p = .085$, but not for unpracticed near-doubles, $F(3.28, 100.08) = .18$, $p = .462$. As the control group was higher on the practiced near-doubles at a marginally significant level at pretest, $F(1, 62) = 2.84$, $p = .097$ (two-tailed), the immediate and delayed posttest comparisons were conducted controlling for initial pretest ability. Effect size calculations for these analyses utilize adjusted means. Contrasts indicated that the intervention group outperformed the control group on immediate posttest at a marginally significant level, $F(1, 63) = 2.67$, $p = .054$, $d = .36$, and significantly outperformed the control group at the delayed posttest, $F(1, 63) = 4.90$, $p = .016$, $d = .50$.

## Question 4: Implications of a Hypothetical Progression of Mental Addition

The fourth set of analyses served to gauge the subsidiary aim of whether the impact of training was affected by hypothesized developmental prerequisites (Implication 1 of the proposed learning progression); a fifth set, whether learning an add-1 rule would have a broader impact (Implication 2).

*Implication 1: Developmental prerequisites specific to a combination family.* Post hoc analyses were conducted to gauge why participants in the near-doubles condition, unlike those in the add-1 conditions, did not profit from intervention. As Frame A of Table 4 shows, all but one child in the add-1 conditions (and one in the near-doubles group) had mastered the developmental prerequisite for efficiently using the add-1 rule with the practiced items 1 + 3, 3 + 1, 1 + 4, 4 + 1, 1 + 7, and 7 + 1 (fluent knowledge of the number-after 3, 4, and 7), and all participants had mastered the prerequisite knowledge for the unpracticed items 1 + 9 and 9 + 1 (fluent knowledge of the number-after 9). In contrast, as Frame B of Table 4 shows, most near-doubles participants (and many or most add-1 children) were not fluent with a developmental prerequisite for efficiently using the near-doubles reasoning strategy with the practiced items 2 + 3, 3 + 2, 3 + 4, 4 + 3 or the unpracticed items 4 + 5 and 5 + 4, namely, the related double or its sum +1 (2 + 2 or 4 + 1, 3 + 3 or 6 + 1, and 4 + 4 or 8 + 1). Note that despite this, five near-doubles participants managed to become fluent with 2 + 3, and one became fluent with four other near-doubles—probably by memorizing these items by rote. In comparison, children in the add-1 conditions who were fluent on a near-double were—to a significant degree—fluent on both prerequisites, rather than not. This result is consistent with these children having invented the near-doubles strategy or having learned it from class instruction or outside school.

*Implication 2: Effects* n *+ 1 or 1 +* n *fluency on learning filler items.* Two sets of analyses evaluated whether children who received add-1 training performed significantly better on practiced and unpracticed filler combinations on each posttest than the children who received near-doubles training. Mauchley's test indicated that the assumption of sphericity was met for practiced items but was marginally violated for unpracticed items, $\chi^2(2) = 5.95$, $p = .052$. In the latter case, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon = .92$). Main effects of time indicated that children made significant improvements in their fluency with practiced filler items, $F(2, 122) = 21.08$, $p < .001$, and unpracticed filler items, $F(1.84, 113.46) = 12.24$, $p < .001$. Interaction contrasts indicated participants who received add-1 training improved significantly more in their fluency than those in the control group on both, $F(4, 120) = 2.04$, $p = .013$ and $F(1.84, 113.46) = 2.04$, $p = .050$. Although children did not perform significantly different on practiced filler items at pretest, $F(1, 61) = .98$, $p = .327$ (two-tailed), participants who received add-1 training performed significantly better than those in the control group at the immediate posttest, $F(1, 61) = 9.83$, $p = .002$, $d = 1.23$, and the delayed posttest, $F(1, 61) = 3.82$, $p = .028$, $d = .87$. Although there was not a significant difference at the pretest, $F(0.92, 56.12) = 2.05$, $p = .157$ (two-tailed), or at the immediate posttest, $F(0.92, 56.12) = 1.26$, $p = .133$, $d = .63$, add-1 participants performed significantly better than control participants on unpracticed filler items at the delayed posttest, $F(0.92, 56.12) = 4.70$, $p = .017$, $d = .94$. All effects from these analyses remained significant after the application of the Benjamini-Hochberg correction.

*Table 4*

**Relation Between Performance on Hypothesized Developmental Prerequisites and Mental Addition Posttest Fluency of *n* + 1/1 + *n* Items (Frame A) and Near Doubles (Frame B) by Condition (Add-1 Conditions vs. Near-Doubles Condition)**

Frame A: Number-after items on the TEMA-3 posttest x add-1 items

| | | | | |
|---|---|---|---|---|
| | 1 + 3 | | 3 + 1 | |
| NA 3 | Not | Fluent | Not | Fluent |
| Fluent | 14 (13) | 29 (7) | 6 (10) | 37 (10) |
| Not | 0 (0) | 0 (1) | 0 (1) | 0 (0) |
| | $p <. 001$ ($p = .001$) | | $p = .016$ ($p = .001$) | |
| | 1 + 4 | | 4 + 1 | |
| NA 4 | Not | Fluent | Not | Fluent |
| Fluent | 11 (6) | 30 (14) | 9 (11) | 32 (9) |
| Not | 2 (1) | 0 (0) | 1 (1) | 1 (0) |
| | $p <. 001$ ($p = .016$) | | $p = .011$ ($p <. 001$) | |
| | 1 + 7 | | 7 + 1 | |
| NA7 | Not | Fluent | Not | Fluent |
| Fluent | 14 (10) | 28 (9) | 12 (9) | 30 (10) |
| Not | 1 (2) | 0 (0) | 1 (2) | 0 (0) |
| | $p <. 001$ ($p = .001$) | | $p <. 001$ ($p = .002$) | |
| | 1 + 6 | | 6 + 1 | |
| NA6 | Not | Fluent | Not | Fluent |
| Fluent | 14 (10) | 14 (10) | 14 (11) | 29 (9) |
| Not | 0 (0) | 0 (0) | 1 (1) | 0 (0) |
| | $p <. 001$ ($p = .006$) | | $p <. 001$ ($p <. 001$) | |
| | 1 + 9 | | 9 + 1 | |
| NA9 | Not | Fluent | Not | Fluent |
| Fluent | 12 (10) | 30 (11) | 11 (12) | 31 (9) |
| Not | 1 (0) | 0 (0) | 1 (0) | 0 (0) |
| | $p <. 001$ ($p = .001$) | | $p <. 001$ ($p <. 001$) | |

Frame B: Doubles and *n*+1 x near doubles

| | | | | |
|---|---|---|---|---|
| | 2 + 3 | | 3 + 2 | |
| 2 + 2 / 4 + 1 | Not | Fluent | Not | Fluent |
| Fluent | 20 (4) | 9 (4) | 22 (4) | 7 (4) |
| Not | 13 (8) | 1 (5) | 13 (12) | 1 (1) |
| | $p = .001$ ($p = .500$) | | $p <. 001$ ($p = .188$) | |
| | 3 + 4 | | 4 + 3 | |
| 3 + 3 / 6 + 1 | Not | Fluent | Not | Fluent |
| Fluent | 19 (6) | 2 (2) | 19 (6) | 2 (2) |
| Not | 20 (12) | 2 (1) | 20 (12) | 2 (1) |
| | $p <. 001$ ($p = .063$) | | $p <. 001$ ($p = .063$) | |

*(continued)*

*Table 4* **(continued)**

| Frame B: Doubles and $n$+1 x near doubles | | | | |
|---|---|---|---|---|
| | 4 + 5 | | 5 + 4 | |
| 4 + 4 / 8 + 1 | Not | Fluent | Not | Fluent |
| Fluent | 18 (5) | 5 (3) | 15 (5) | 8 (3) |
| Not | 19 (13) | 1 (0) | 20 (12) | 0 (1) |
| | $p$ <. 001 ($p$ = .031 ) | | $p$ <. 001 ($p$ = .109) | |

*Note*. Data without parentheses are for the add-1 conditions; data in parentheses are for the near-doubles condition. For Frame A, the developmental prerequisite for efficient use of the add-1 rule for a 1 + 3 or a 3 + 1, for example, is fluency with the number-after relation 3 is 4 (NA 3). For Frame B, needed for efficient use of the near-doubles strategy for 2 + 3 and 3 + 2, for instance, are fluency with the double 2 + 2 and 4 + 1. "Fluent" for the developmental prerequisites indicates that a child was fluent on both the related double and n + 1 item; the "Not" indicates that a child was not fluent on one or both prerequisites. A significant p value (test for the equality of partially overlapping frequency) confirms that a hypothesized prerequisite is a necessary condition for fluency with an item.

## Discussion

The results regarding the four questions are discussed in turn.

### Question 1: Feasibility of the Add-1 Training

The significant results for the practiced $n$ + 1 or 1 + $n$ items and, particularly, the effect sizes indicative of effective practice for the unpracticed $n$ + 1 or 1 + $n$ items at both posttests indicate that computer-assisted discovery learning is feasible for helping first graders with a risk factor learn the add-1 rule. Specifically, the add-1 training—which entailed minimal extra practice of $n$ + 1 or 1 + $n$ items and number-after relations (only 10 times each), feedback on correctness only, and implicit pattern detection (with minimal scaffolding or not)—enabled participants to construct *and* retain a general add-1 rule and flexibly/appropriately apply it to new items. The results were achieved *above and beyond* those produced by regular classroom instruction and practice that focused on learning $n$ + 1 or 1 + $n$ items (see Table 2), including that in Schools 2 to 4, which used *Everyday Mathematics*. Taken together, these results indicate that the effects of the add-1 interventions were highly robust.

Unlike children in the comparison group—including those who used the *Everyday Mathematics* program—practicing number-after relations and $n$ + 1 sums successively (as in the structured add-1 condition) or even during the same session (as in the unstructured add-1 condition)—if only 10 times each—may have increased the chances that add-1 participants considered both simultaneously in working memory, thus permitting the insight that is the add-1 rule. This conclusion is an extension of the proposal that practice is effective in strengthening a specific association only when an expression and its answer are simultaneously present in working memory (Fuchs et al., 2006; Geary, 1993). Moreover, the appreciable effect sizes when children who had the supplemental add-1 training plus *Everyday Mathematics* (UCSMP, 2005) are compared to peers who had only *Everyday Mathematics* is consistent with the generation effect—the enhancement of learning and retention when learners are encouraged to generate their own generalization. That is, the supplemental add-1 discovery learning produced better learning of the add-1 rule than direct instruction alone on the topic.

Moreover, as Figures 5 and 6 show, the add-1 training most helped children with the lower pretest fluency. This is an encouraging finding for developing early intervention for children with a risk factor and may help to explain why including or excluding the two high achievement participants from the structured add-1 condition did not make a difference.

Even so, the add-1 training undoubtedly could be improved. The fact that more number-after-$n$ (and other) response biases were maintained or emerged for the add-1 group than for the comparison group indicates the need to modify add-1 programs so that they foster flexible and appropriate application of number-after knowledge to addition. Intermixing nonexamples of the relation (sometimes posing, e.g., $4 + 2 = ?$ immediately after the question "What number comes number after 4?") would require children who to apply their knowledge of number-after relations more thoughtfully.[3] Although confusing for at-risk preschoolers (Baroody, Eiland, et al., 2009), including subtraction items such as $4 - 1$ for at-risk first graders might provide an even sharper contrast and prompt flexible use of number-after and number-before knowledge.

### Question 2: Comparison of Guided and Unguided Discovery

Contrary to Alfieri et al. (2011) and Kirschner et al.'s (2006) conclusions, participants who received unstructured add-1 training achieved comparable gains in fluency with $n + 1$ or $1 + n$ items as those who received the structured add-1 training. Although this pattern of results suggests that additional practice is all that is needed to promote fluency with these basic combinations, several considerations suggest that participants in both the guided *and* unguided add-1 training probably discovered the add-1 rule rather than memorized $n + 1$ or $1 + n$ facts by rote. First, these results were achieved *despite* only 10 repetitions for each of the eight practiced $n + 1$ or $1 + n$ items (and no repetitions for transfer items). This is substantially less practice than thousands of repetitions per item necessary to achieve memorization (by rote) of these facts specified by earlier models and computer simulations of arithmetic learning (e.g., Shrager & Siegler, 1998;

Siegler & Jenkins, 1989).[4] Second, transfer to unpracticed $n + 1$ or $1 + n$ items is consistent with rule-governed, not rote, learning.

Why were the structured and unstructured add-1 training equally effective in helping participants induce the add-1 rule? One—and the most likely—explanation is that the link between number-after-$n$ relations and adding 1 is relatively salient. The participants in both add-1 conditions had fluent knowledge of number-after-$n$ relations, and the connection between this knowledge and adding 1 is relatively obvious. This may explain Alfieri et al.'s (2011) often-overlooked finding that their outcomes were moderated by domain and age—that the magnitude of effects were smaller for certain academic domains (notably mathematics) and for different ages (notably for younger children). Relatively speaking, the domain of early childhood mathematics offers many relatively salient patterns and relations that children can discover without much guidance, whereas other domains and higher-level mathematics involve patterns and relations that are more abstruse and require more guidance to notice.

A second seemingly plausible explanation for the similar results of the unguided and guided discovery in the present study—as opposed to the significant differences found by Alfieri et al. (2011)—is that the differences between unstructured and structured add-1 training were relatively small. The only differences were that the structured condition involved solving $n + 1$ items immediately after answering the related number-after-$n$ item and the juxtaposition of the two equations in the feedback of add-1 sums. Although these features by themselves were apparently ineffective, they could easily be supplemented to better highlight the connection between the number-after relations and adding 1 in two ways:

- Games with number lists have been shown to improve performance on a variety of number tasks (Ramani & Siegler, 2008; Siegler & Ramani, 2009) and might be particularly useful in discovering the add-1 rule. Specifically, a number list could be included below the columns of the Castle Wall and items would be ordered by the size of their answer instead of by the order a set of items is presented (e.g., the fourth item practiced, $1 + 4 = 5$, would settle in the fifth column labeled 5, not the fourth unnumbered column as in shown Frame D of Figure 2).
- To better ensure the simultaneous representation of a number-after relation and a related $n + 1$ or $1 + n$ combination in working memory and, thus, the opportunity to discover the connection between these aspects of knowledge, attention could be explicitly drawn to this connection by, for instance, a "Does It Help?" activity. This would entail asking a child to consider whether knowing that the number after five is six would help answer $5 + 1$ and providing implicit feedback ("Yes, knowing that the number after five is six can help answer $5 + 1$") or explicit feedback ("Yes, because the answer to $5 + 1$ is the number after five when counting, which is six").

However, subsequent research that included such additional scaffolding in the structured add-1 training also failed to produce significantly better learning or transfer than unstructured add-1 training (Purpura et al., 2012).

Whether minimally guided and implicit discovery of the add-1 rule might be more effective than essentially unguided discovery for younger, less developmentally advanced children needs to be examined. Future research also needs to explore whether different versions of guided discovery might differ in effectiveness. “Discovery learning [can] range from implicit pattern detection [used in both the unstructured and structured add-1 conditions in the present study] . . . to the elicitation of explanations” (Alfieri et al., 2011, p. 2). The structured program might be improved by supplementing active and constructive activities with interactive ones, such as dialoguing with a virtual character, considering its contribution, and revising errors based on scaffolding or feedback (Chi, 2009). Eliciting explanations has also been found to be an effective way to foster strategy learning and generalization (Chi, de Leeuw, Chiu, & LaVancher, 1994; Crowley & Siegler, 1999). See Hmelo-Silver et al. (2007) for additional forms of scaffolding.

### Question 3: Feasibility of Near-Doubles Training

The effect sizes, which met criterion for effective practice set by the federal What Works Clearinghouse (IES, 2011), indicated that the computer-assisted, structured near-doubles training for first graders with at least one risk factor was feasible in prompting lasting learning of the practiced, but not the unpracticed, near-doubles. Whether the near-doubles training can be improved or such training should be postponed until second grade needs further study.

### Question 4: Implications of the Proposed Trajectory of Mental Addition Development

*Implication 1: Developmental prerequisites specific to a combination family.* The differential success of the add-1 and near-doubles training paralleled different levels of fluency with developmental prerequisites specific to the target family. With the add-1 conditions, nearly all participants had mastered the number-after relations necessary for learning and efficiently using the add-1 rule. In contrast, most participants in the near-doubles condition were not fluent with the double and/or $n + 1$ or $1 + n$ items necessary to learn or efficiently use the near-doubles strategy. As Frame B in Table 4 shows, the single substantial exception was that five near-doubles participants mastered the practiced combination $2 + 3$ but not one of both developmental prerequisites ($2 + 2 = 4$ and $4 + 1 = 5$). Given the lack of transfer to the unpracticed but related item $3 + 2$, it appears that four of these five exceptional cases were the result of memorizing $2 + 3 = 5$ by rote with minimal but focused additional practice.

*Implication 2: Effects* n + *1 or 1* + n *fluency on learning filler items.* Participants in the add-1 conditions may have been more successful than those in the near-doubles condition in mastering the filler items practiced by all (3 + 5 and 5 + 7) and the filler's unpracticed commuted counterparts (5 + 3 and 7 + 5), because nearly all of the combinations they encountered during the training could be answered by using relatively salient add-0 and add-1 rules. These children, then, had to memorize only two combinations (3 + 5 and 5 + 7) by "brute force," *and* perhaps had relatively fluent 1 + $n$ sums (particularly 3 + 1 = 4 and 1 + 7 = 8) as a reference point. As commutativity is readily apparent to most first graders, this fluency was transferred to nonpracticed items 5 + 3 and 7 + 5. For near-doubles participants, 3 + 5 and 5 + 7 were just two of many combinations they needed to memorize by brute force.

## Conclusions

In conclusion, the present research indicates that computer-assisted instruction is a feasible means of helping first graders with a risk factor discover the add-1 rule via minimally guided *or* essentially unguided implicit pattern detection (cf. Alfieri et al., 2011; Kirschner et al., 2006). Those who began the second semester of first grade with little or no fluency with $n$ + 1 or 1 + $n$ items, in particular, benefitted from the add-1 training. The effectiveness of the add-1 training in promoting fluency with unpracticed $n$ + 1 or 1 + $n$ items and practiced and unpracticed filler items such as 3 + 5 and 7 + 3 provides additional supporting evidence for the generation effect with a genuine school (ecologically valid) task. It also corroborates positions outlined by both the NMAP (2008) and the number sense view that *both* forming associations via practice and recognizing relations are key aspects of achieving fluency and that reasoning processes can be an efficient basis for determining the solutions to basic combinations.

The results of the present study are also consonant with the proposed hypothetical learning trajectory that add-1 rule is the connection between the representations of counting and numerical magnitude (the successor principle/integer representation/linear representation) and retrieval structures. Information stored in long-term memory may not have a permanent form and may be changed each time the memory is recalled (Nader & Hardt, 2009). Practice, then, is not merely a vehicle for strengthening a factual association but an opportunity to enrich memory of a combination by actively creating new connections with it. For example, recalling that the number after "seven" is "eight" while solving (calculating) 7 + 1 = ? may help children to construct or strengthen the successor principle (each counting number is exactly one more than its predecessor), which is arguably the conceptual basis for a linear representation of numbers. Recognizing the connection between adding 1 and known number-after relations appears to quickly translate into the generalization that the sum of *any* $n$ + 1 item is the successor of $n$. Such a representation allows children to use their (automatic) knowledge of the generative rules for counting to efficiently deduce the sum of any $n + 1$ or $1 + n$ item for any known part of the counting sequence. For all of these reasons, learning the add-1 rule should be a focal point or primary goal of first-grade instruction for all pupils.

Computer-assisted instruction involving minimal scaffolding can help first graders achieve fluency with the easiest sums with one caveat. A child must have the developmental readiness or prerequisites for the reasoning strategy. The implementation of the near-doubles training in the present study mimics the all too common practices of using (a) lockstep instruction—moving on to a new unit before all children have had the opportunity to master previous units in an effort to get through a curriculum—and (b) premature drill—practice with combinations before a child understands the underlying relation(s) of a combination family (Brownell & Chazal, 1935; NRC, 2001; Rathmell, 1978).

Disquieting is that with regular classroom instruction and practice, the near-doubles participants were fluent on less than half of the relatively easy $n + 1$ or $1 + n$ combinations at the *end* of first grade, and all participants had mastered less than two-thirds of doubles. These results have the following educational implications: (a) Educators cannot take for granted that first graders with a risk factor will become fluent with even the easiest sums. (b) As meaningful learning of basic combinations entails building on existing or familiar knowledge, educators need to consider carefully whether a child has mastered the developmental prerequisites for a particular combination family, as well as those for mental addition in general. Diagnostic information on such prerequisites is essential for successful differentiated instruction. (c) Practice is an instructional tool that needs to be used judiciously. Specifically, practicing a combination family before mastering its developmental prerequisites may be ineffective in promoting meaningful learning. Also, practice should be designed to prompt the simultaneous consideration of different aspects of relational knowledge in working memory. (d) There is a need for educational software to supplement regular classroom instruction on basic combinations—software that provides simultaneous or proximate practice of number-after relations and adding 1 and doubles and everyday visual analogies. Stand-alone software could provide cost-effective training that many teachers do not have the time and training to deliver.

## Notes

The research reported herein was supported by a grant from the Institute of Education Science, U.S. Department of Education, through Grant R305K050082 ("Developing an Intervention to Foster Early Number Sense and Skill"). Preparation of this manuscript also was supported, in part, by grants from the U.S. Department of Education (R305A080479: "Fostering Fluency With Basic Addition and Subtraction

Facts") and (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 Department of Education.

[1]The doubles are relatively easy to learn because they embody familiar real-world pairs of a set, such as 5 fingers on one hand and 5 on the other is 10 fingers altogether and one row of 6 eggs and a second row is a dozen or 12 eggs altogether (Baroody, 1989a; Rathmell, 1978). Moreover, the sums of doubles are even numbers (which parallel the skip-count-by-two sequence: "2, 4, 6 . . . ") and are akin to the first two counts in various skip counts (e.g., 5 + 5 can be reinforced by knowing the skip-count-by fives: "Five, ten")—common aspects of primary-level mathematics instruction (Baroody, 1998).

[2]Although it is not possible to distinguish between using the add-1 rule and abstract counting-on, research by the first author (Baroody, 1995) and Bråten (1996) indicate that the former develops prior to and provides a basis for inventing the latter. Specifically, within days of discovering the add-1 rule, children begin to counting-on two and then three (e.g., "5 + 2, one more is six, two more is seven"). The add-1 rule is not necessary for learning the less sophisticated concrete counting-on strategy (e.g., for 5 + 2: put up two fingers, state the cardinal value of the larger addend "five" and count from there on the fingers "six" [pointing to the first finger], *seven* [pointing to the second finger]). However, children would have to represent an addend ahead of time. This takes time and would probably be seen by a trained observer.

Research indicates that some mental addition novices are prone to using response bias that indiscriminately, and in some cases can lead to false positives (Baroody, 1989b, 1992; Baroody, Purpura, Reid, & Eiland, 2011; Dowker, 1997, 2003). For example, the common biases of "stating an addend" and "state the number after the larger addend" accidentally results in correct answers to $n + 0$ or $0 + n$ items and $n + 1$ or $1 + n$ items, respectively.

[3]At pretest, three children in each condition exhibited a number-after-$n$ or number-after-larger-$n$ response bias, and two more comparison children exhibited other mechanical responses biases. At the delayed posttest, four in the add-1 treatments still stated a number-after an addend unselectively, and two other children adopted this or another response bias. In contrast, only one child in the comparison group exhibited a response bias.

[4]For $n + 1$ items, the total number repetitions would be 30 if practice with the related number-after-$n$ item and—as suggested by some (but not all) models of mental addition development—practice of a commuted $1 + n$ partner were counted. This is still relatively little practice.

## References

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*, 1–18.

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

Baratta-Lorton, M. (1995). *Mathematics their way*. Menlo Park, CA: Addison-Wesley.

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

Baroody, A. J. (1987). *Children's mathematical thinking: A developmental framework for preschool, primary, and special education teachers*. New York, NY: Teachers College Press.

Baroody, A. J. (1989a). *A guide to teaching mathematics in the primary grades*. Boston, MA: Allyn & Bacon.

Baroody, A. J. (1989b). Kindergartners' mental addition with single-digit combinations. *Journal for Research in Mathematics Education*, *20*, 159–172.
Baroody, A. J. (1992). The development of kindergartners' mental-addition strategies. *Learning and Individual Differences*, *4*, 215–235.
Baroody, A. J. (1995). The role of the number-after rule in the invention of computational short cuts. *Cognition and Instruction*, *13*, 189–219.
Baroody, A. J. (with Coslick, R. T.). (1998). *Fostering children's mathematical power: An investigative approach to K–8 mathematics instruction*. Mahwah, NJ: Erlbaum.
Baroody, A. J., Bajwa, N. P., & Eiland, M. D. (2009). Why can't Johnny remember the basic facts? *Developmental Disabilities Research Reviews*, *15*, 69–79.
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.
Baroody, A. J., Eiland, M. D., & Thompson, B. (2009). Fostering at-risk preschoolers' number sense. *Early Education and Development*, *20*, 80–120.
Baroody, A. J., Purpura, D. J., & Reid, E. E. (2012). Comments on learning and teaching early and elementary mathematics. In J. Carlson & J. Levin (Eds.), *Instructional strategies for improving students' learning* (pp. 163–175). Charlotte, NC: Information Age Publishing.
Baroody, A. J., Purpura, D. J., Reid, E. E., & Eiland, M. D. (2011, September). Scoring fluency with basic addition combinations in context. In Baroody, A. J. (Chair), *Issues in the assessment and scoring of early numeracy skills*. Symposium conducted at the annual meeting of the Society for Research on Educational Effectiveness, Washington, DC.
Baroody, A. J., & Varma, S. (in press). The active construction view of basic number fact knowledge: New directions for cognitive neuroscience. In J. Baek, A. E. Kelly, & L. Kalbfleisch (Eds.), *Neuropsychology and mathematics education*.
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.
Bråten, I. (1996). *Cognitive strategies in mathematics* (Report No. 10). Oslo, Norway: Institute for Educational Research, University of Oslo.
Brownell, W. A. (1935). Psychological considerations in the learning and the teaching of arithmetic. In W. D. Reeve (Ed.), *The teaching of arithmetic* (pp. 1–31). New York, NY: Bureau of Publications, 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.
Brownell, W. A., & Chazal, C. B. (1935). The effects of premature drill in third-grade arithmetic. *Journal of Educational Research*, *29*, 17–28.
Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, *31*, 21–32.
Canobi, K. H., Reeve, R. A., & Pattison, P. E. (1998). The role of conceptual understanding in children's problem solving. *Developmental Psychology*, *34*, 882–891.
Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*, 73–105.
Chi, M. T. H., de Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*, 439–477.
Clements, D. H., & Sarama, J. (Eds.). (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning*, *6*, 81–89.
Clements, D. H., & Sarama, J. (2009). *Learning and teaching early math: The learning trajectories approach*. New York, NY: Routledge.
Clements, D. H., & Sarama, J. (2012). Learning and teaching early and elementary mathematics. In J. Carlson & J. Levin (Eds.), *Psychological perspectives on*

*contemporary educational issues: Instructional strategies for improving students' learning* (pp. 107–162). Charlotte, NC: Information Age Publishing.

Cohen J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159.

Common Core State Standards. (2011). *Common Core State Standards: Preparing America's students for college and career.* Retrieved from http://www.carnegielearning.com/webinars/common-core-standards-initiative-preparing-americas-students-for-college-and-career/.

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: 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.

Crowley, K., & Siegler, R. S. (1999). Explanation and generalization in young children's strategy learning. *Child Development*, *70*, 304–316.

Dev, O. C., Doyle, B. A., & Valente, B. (2002). Labels needn't stick: "At-risk" first graders rescued with appropriate intervention. *Journal of Education for Students Placed At Risk*, *7*, 327–332.

Dowker, A. D. (1997). Young children's addition estimates. *Mathematical Cognition*, *3*, 141–154.

Dowker, A. D. (2003). Young children's estimates for addition: The zone of partial knowledge and understanding. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 35–74). Mahwah, NJ: Erlbaum.

Dowker, A. (2009). Use of derived fact strategies by children with mathematical difficulties. *Cognitive Development*, *24*, 401–410.

Fayol, M., & Thevenot, C. (2012). The use of procedural knowledge in simple addition and subtraction. *Cognition*, *123*, 392–403.

Ferguson, R. F. (1998). Can schools narrow the Black-White test score gap? In C. Jencks & M. Phillips (Eds.), *The Black-White test score gap* (pp. 318–374). Washington, DC: Brookings Institution Press.

Fuchs, L. S., Compton, D., Fuchs, D., Paulsen, K., Bryant, J., & Hamlett, C. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology*, *97*, 493–513.

Fuchs, L. S., Fuchs, D., Hamlet, C. L., Powell, S. R., Capizzi, A. M., & Seethaler, P. M. (2006). The effects of computer-assisted instruction on number combination skill in at-risk first graders. *Journal of Learning Disabilities*, *39*, 467–475.

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., & Hamlett, C. L. (2010). The effects of strategic counting instruction, with and without deliberate practice, on number combination skill among students with mathematics difficulties. *Learning and Individual Differences*, *20*, 89–100.

Fuson, K. C. (1988). *Children's counting and concepts of number.* New York, NY: Springer-Verlag.

Fuson, K. C. (1992). Research on whole number addition and subtraction. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 318–374). New York, NY: Macmillan.

Geary, D. C. (1990). A componential analysis of an early learning deficit in mathematics. *Journal of Experimental Child Psychology*, *49*, 363–383.

Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin*, *114*, 345–362.

Geary, D. C. (1996). *Children's mathematical development: Research and practical applications*. Washington, DC: American Psychological Association.
Geary, D. C. (2003). Arithmetic development: Commentary on chapters 9 through 15 and future directions. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 453–464). Mahwah, NJ: Erlbaum.
Gersten, R., & Chard, D. (1999). Number sense: Rethinking arithmetic instruction for students with mathematical disabilities. *The Journal of Special Education*, *33*, 18–28.
Ginsburg, H. P., & Baroody, A. J. (2003). *Test of early mathematics ability* (3rd ed.). Austin, TX: Pro-Ed.
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. *Journal for Research in Mathematics Education*, *39*, 153–183.
Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark. *Educational Psychologist*, *42*, 99–107.
Institute of Education Sciences. (2011). *What works clearinghouse: Procedures and standards handbook (Version 2.1)*. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/reference_resources/
wwc_procedures_v2_1_standards_handbook.pdf.
James, W. (1958). *Talks to teachers on psychology: And to students on some of life's ideals*. New York, NY: W. W. Norton & Company. (Talk originally given 1892)
Jerman, M. (1970). Some strategies for solving simple multiplication combinations. *Journal for Research in Mathematics Education*, *1*, 95–128.
Jordan, N. C. (2007). The need for number sense. *Educational Leadership*, *65*, 63–66.
Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003). Arithmetic fact mastery in young children: A longitudinal investigation. *Journal of Experimental Child Psychology*, *85*, 103–119.
Jordan, N. C., Hanich, L. B., & Uberti, H. Z. (2003). Mathematical thinking and learning difficulties. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 359–383). Mahwah, NJ: Erlbaum.
Jordan, N. C., Huttenlocher, J., & Levine, S. C. (1992). Differential calculation abilities in young children from middle- and low-income families. *Developmental Psychology*, *28*, 644–653.
Jordan, N. C., Huttenlocher, J., & Levine, S. C. (1994). Assessing early arithmetic abilities: Effects of verbal and nonverbal response types on the calculation performance of middle- and low-income children. *Learning and Individual Differences*, *6*, 413–432.
Jordan, N. C., Kaplan, D., Olah, L. N., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development*, *77*, 153–175.
Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, *45*, 850–867.
Jordan, N. C., & Levine, S. C. (2009). Socioeconomic variation, number competence, and mathematics learning difficulties in young children. *Developmental Disabilities Research Reviews*, *15*, 60–68.

Katona, G. (1967). *Organizing and memorizing: Studies in the psychology of learning and teaching*. New York: Hafner. (Original work published 1940)

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*, 75–86.

Knotternerus, J. A., & Bouter, L. M. (2001). The ethics of sample size: Two-sided testing and one-sided thinking. *Journal of Clinical Epidemiology*, *54*, 109–110.

Lee, J. (2004). Multiple facets of inequality in racial and ethnic achievement gaps. *Peabody Journal of Education*, *79*, 51–73.

Lubienski, S. T. (2001, April). *Are the NCTM standards reaching all students? An examination of race, class, and instructional practices*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Lubienski, S. T. (2002). A closer look at Black-White mathematics gaps: Intersections of race and SES in NAEP achievement and instructional practices data. *Journal of Negro Education*, *71*, 269–287.

Lubienski, S. T., & Shelley II, M. C. (2003, April). *A closer look at U.S. mathematics instruction and achievement: Examinations of race and SES in a decade of NAEP data*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ: Erlbaum.

Mazzocco, M., & Thompson, R. (2005). Kindergarten predictors of math learning disability. *Learning Disabilities Research & Practice*, *20*, 142–155.

Miller, D., & McKinnon, A. (1995). *Beginning School Mathematics Project: An ASCD curriculum handbook case study section 11H*. Alexandria, VA: Association for Supervision and Curriculum Development.

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.

Nader, K., & Hardt, O. (2009). A single standard for memory: The case for reconsolidation. *Nature Reviews/Neuroscience*, *10*, 224–234.

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.

National Research Council. (2009). *Mathematics in early childhood: Learning paths toward excellence and equity*. Washington, DC: National Academy Press.

Piaget, J. (1964). Development and learning. In R. E. Ripple & V. N. Rockcastle (Eds.), *Piaget rediscovered* (pp. 7–20). Ithaca, NY: Cornell University.

Pillemer, D. B. (1991). One- versus two-tailed hypothesis tests in contemporary educational research. *Educational Researcher*, *20*, 13–17.

Purpura, D. J., Baroody, A. J., Eiland, M. D., & Reid, E. E. (2012, September). Fostering first-graders' reasoning strategies with the most basic sums. In Baroody, A. J. (Chair), *Implications of three training experiments for tailoring instruction to promote the fluency of specific arithmetic families*. Symposium

at the annual meeting of the Society for Research on Educational Effectiveness, Washington, DC.
Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development*, *79*, 375–394.
Rasanen, P., Salminen, J., Wilson, A., Aunio, P., & Dehaene, S. (2009). Computer-assisted intervention for children with low numeracy skills. *Cognitive Development*, *24*, 450–472.
Rathmell, E. C. (1978). Using thinking strategies to teach basic facts. In M. N. Suydam & R. E. Reys (Eds.), *Developing computational skills* (pp. 13–50). Reston, VA: National Council of Teachers of Mathematics.
Resnick, L. B. (1983). A developmental theory of number understanding. In H. P. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 109–151). New York, NY: Academic Press.
Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research: Learning trajectories for young children*. New York, NY: Routledge.
Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science*, *9*, 405–410.
Siegler, R. S. (1986). Unities across domains in children's strategy choices. In M. Perlmutter (Ed.), *Minnesota symposium of child psychology* (*Vol. 19*, pp. 1–48). Hillsdale, NJ: Erlbaum.
Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, NJ: Erlbaum.
Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology*, *101*, 545–560.
Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 592–604.
Steinberg, R. M. (1985). Instruction on derived facts strategies in addition and subtraction. *Journal for Research in Mathematics Education*, *16*, 337–355.
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.
Swanson, H. L., & Cooney, J. (1985). Strategy transformation in learning disabled and nondisabled students. *Learning Disability Quarterly*, *8*, 221–230.
Swanson, H. L., & Rhine, B. (1985). Strategy transformation in learning disabled children's math performance: Clues to the development of expertise. *Journal of Learning Disabilities*, *18*, 596–603.
Swenson, E. J. (1949). Organization and generalization as factors in learning, transfer and retroactive inhibition. In *Learning theory in school situations*. Minneapolis, MN: University of Minnesota Press.
Thiele, C. L. (1938). *The contribution of generalization to the learning of addition facts*. New York, NY: Bureau of Publications, Teachers College, Columbia University.
Thorndike, E. L. (1922). *The psychology of arithmetic*. New York, NY: Macmillan.
Thornton, C. A. (1978). Emphasizing thinking strategies in basic fact instruction. *Journal for Research in Mathematics Education*, *9*, 214–227.
Thornton, C. A. (1990). Solution strategies: Subtraction number facts. *Educational Studies in Mathematics*, *21*, 241–263.
Thornton, C. A., & Smith, P. J. (1988). Action research: Strategies for learning subtraction facts. *Arithmetic Teacher*, *35*, 8–12.

Torbeyns, J., Verschaffel, L., & Ghesquiere, P. (2005). Simple addition strategies in a first-grade class with multiple strategy instruction. *Cognition and Instruction*, *23*, 1–21.

Torgeson, J. K., & Young, K. A. (1983). Priorities for the use of microcomputers with learning disabled children. *Journal of Learning Disabilities*, *16*, 234–237.

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.

Wilburn, D. B. (1949). *Methods of self instruction for learning arithmetic*. Chicago: University of Chicago.

Wilkinson, L., & APA Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*, 594–604.

Wynroth, L. (1986). *Wynroth math program—The natural numbers sequence*. Ithaca, NY: Wynroth Math Program. (Original work published 1975)

Manuscript received November 4, 2011
Final revision received October 10, 2012
Accepted November 27, 2012