



# Intensifying Instruction to Address Treatment Resistance to Early Mathematics Interventions

Maryam Nozari<sup>1</sup> · Diane P. Bryant<sup>2</sup> · Christian T. Doabler<sup>2</sup> · Terry S. Falcomata<sup>2</sup> · Meijia Liu<sup>2</sup> · Rene Grimes<sup>3</sup>

Accepted: 15 September 2021 / Published online: 29 September 2021

This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2021

## Abstract

Previous studies demonstrate the positive effects of explicit mathematics intervention on the mathematics outcomes of students at risk for mathematics difficulties (MD). One such intervention was the Tier 2 Booster Intervention (T2BI), a supplemental mathematics intervention focused on whole number concepts and skills that detailed the explicit and systematic instructional design of T2BI. However, findings from past research suggest that some students do not adequately respond to this purposefully designed intervention. Against that backdrop, the focus of this study is to address this notion of minimal response to intervention by integrating a select set of validated Cognitive Learning Principles (CLPs) into T2BI and investigating the extent to which students with MD respond favorably to the resulting intervention. Employing a multiple baseline design, our results showed a functional relation between the provision of an intervention with CLPs and increasing accuracy scores on both proximal and distal measures across the four participants with MD. These findings and their broader implications using CLPs to reduce minimal response to well-designed mathematics *interventions* are discussed.

**Keywords** Early intervention · Explicit instruction · Cognitive learning principles · Mathematics education · Cognitive psychology

According to the National Assessment of Educational Progress (NAEP, 2019), the achievement gap between students with mathematics difficulties (MD) and typically performing students has increased over the past decade (National Center for Education Statistics, 2019). Fourth-grade students who demonstrate low mathematics achievement continue to have problems in mathematics compared to their typically developing peers (i.e., scoring below the NAEP basic level). As early as kindergarten and first grade, the achievement gaps are present between the average students and students who enter school with a poor understanding of mathematics (Strand Cary et al., 2017). Students with MD in early grades demonstrate more procedural errors and have

persistent problems with quick retrieval of basic facts than their peers (Geary, 2011; Gersten et al., 2005; Shrager & Siegler, 1998). This persistent problem is alarming because students acquire foundational skills in the early grades. It is important to find ways to facilitate conceptual understanding of the fundamental skills for such at-risk students. As such, early mathematics interventions for at-risk students have been developed to shorten the gap between students with MD and their peers (Bryant, 2008; Clarke et al., 2017; Gersten et al., 2015; Jordan et al., 2012; Klein et al., 2008).

Most existing interventions incorporate features associated with positive effects, including explicit instruction, cumulative review, multiple representations, comprehensiveness of taught skills, and progress monitoring (Fuchs et al., 2014; Gersten et al., 2009). However, while the overall effects of the designed interventions were positive for most students with MD (Wang et al., 2016), there is a small sub-sample of students who receive these interventions and are still struggling to demonstrate adequate progress relative to their at-risk peers (Clarke et al., 2014). These students are referred to as low responders (Vaughn et al., 2009) or treatment-resisters (Vellutino et al., 1996).

Low responders (treatment-resisters) show specific characteristics, resulting in lower progress in early interventions (Al

✉ Maryam Nozari  
nozari@hawaii.edu

<sup>1</sup> Department of Special Education, University of Hawaii at Manoa, 1776 University Avenue, Wist Hall Room 120, Honolulu, HI 96822, USA

<sup>2</sup> Department of Special Education, The University of Texas at Austin, Austin, TX, USA

<sup>3</sup> Department of Curriculum and Instruction, Tennessee Technological University, Cookeville, TN, USA

Otaiba & Fuchs, 2002), which leads to a bigger gap between low responders and their average-performing peers as they move through higher grades. Thus, it is important to fast-track these students and provide more intensive individualized instruction in early grades (Compton et al., 2012). Fuchs et al. (2013) suggested that individual differences in cognitive resources such as non-verbal reasoning played an important role in determining responsiveness to the mathematics intervention (Kearns & Fuchs, 2013). For example, students with lower reasoning skills are more likely to respond inadequately to a generally effective intervention (Barner et al., 2016). Low responders may also struggle to notice conceptual similarities across problems, leading to difficulties in generalization and transferring their knowledge to solve new tasks (Morsanyi et al., 2013). Overall, it is important to note that the evidence of domain-general skills in intervention responsiveness is mixed. More research data are needed to specify the potential moderator and mediator effects of both domain-general and domain-specific abilities of low responders for growth in mathematics. Although it is unclear whether the domain-general factors contribute to the response to early mathematics intervention, there is a great deal of evidence showing relations between difficulties in several aspects of mathematical cognition and deficits in these more general cognitive systems (Barnes & Raghobar, 2017). Therefore, it is important to consider the specific characteristics of low responders in designing and implementing early mathematics interventions.

To address the specific characteristics of low responders, researchers have proposed solutions to increase intervention effectiveness for these students (Compton et al., 2012; Fuchs et al., 2014; Vaughn et al., 2009). These solutions include significantly intensifying interventions, increasing the comprehensiveness of taught skills and strategies, explicitly teaching for transfer, and addressing students' cognitive and linguistic limitations with learning disabilities (Fuchs et al., 2014). Early mathematics interventions that incorporated these solutions have been developed and investigated. Clarke et al. (2014) examined the efficacy of a Tier 2 mathematics intervention for first-grade students. The intervention was significantly intensive (e.g., delivered in small-group sessions five days per week) and incorporated the explicit and systematic instructional design principles. However, there was a subgroup of students who did not respond to the intervention. As such, Clarke et al. (2014) recommended finding a set of modifications and systematic manipulation of existing interventions that could improve responsiveness in low responders (Clarke et al., 2014). Another validated early mathematics intervention for low-performing students in first and second grades was Tier 2 Booster Intervention (T2BI) (Bryant et al., 2008). The researchers examined the effect of T2BI on the performance of first and second-grade students who were identified as having MD using a regression discontinuity design. The intervention was provided to low-performing students in a small group,

five days per week for a total of 18 weeks. The regression discontinuity analysis showed an overall moderate positive effect ( $b=0.19$ ,  $p=0.018$ ) on students' mathematics outcomes. However, like other validated interventions for students with MD, a significant subsample of first-grade students showed minimal response to the T2BI ( $b=0.04$ ). The results of a study that examined the effects of T2BI on the performance of first-grade students who were identified as at risk for MD showed that despite the explicit and systematic instructional design of T2BI, some students did not adequately respond to this purposefully designed intervention. Bryant et al. (2008) suggested intensive, relentless, iterative, individualized instruction for these students. The authors also suggested that future studies examine additional tutoring features to help Tier 2 students with low levels (i.e., flat slope) of response. Thus, it is important to use techniques that meet the specific needs of low responders and increase their intervention responsiveness.

Cognitive Learning Principles (CLPs) have promising characteristics which in combination with the existing interventions, can improve responsiveness in low responders. Cognitive psychologists and researchers have identified and developed CLPs that could improve students' learning and lead to generalization and retention of information across mathematics concepts (Booth et al., 2017). CLPs are easy-to-use learning techniques that can be used to improve students' achievement. Many of these CLPs are inexpensive to implement and potentially useful for mathematics instruction (Dunlosky et al., 2013; Roediger & Pyc, 2012). The CLPs are not mutually exclusive categories but include retrieval practice (testing effect); type and timing of feedback; distributed practice, interleaved practice, self-explanation, worked example, multiple representations, analogical comparison, and error reflection (Booth et al., 2017; Dunlosky et al., 2013). Several of these CLPs have been tested in applied learning studies for students across different grade levels and showed promising effects (Dunlosky et al., 2013). Koedinger et al. (2013) indicated that CLPs promote three different functions of instruction, including memory and fluency (e.g., distributed practice and feedback); induction and refinement (e.g., worked example, multiple representations, and interleaving practice); and sense-making and understanding (e.g., error reflection and analogical comparison) (Koedinger et al., 2013). This study aims to examine the provision of an existing intervention with CLPs to improve intervention responsiveness and address the unique learning needs of low responders.

## Rationale for This Study

Although previous research shows that CLPs can improve mathematics instruction, both for simple and more complex skills and concepts, to our knowledge, no study examines

the effects of CLPs for improving response to existing mathematics interventions. The objective of this study is to investigate the effectiveness of the integration of Cognitive Learning Principles (CLPs) into the existing T2BI on improving the intervention effects for low responders. The CLPs that have been integrated into the design of T2BI include self-explanation, interleaving practice, error reflection, and worked example. In summary, the following research questions guided this study: (a) Does integrate CLPs and T2BI result in improved mathematics performance of low responders? (b) Are the effects of the intervention maintained two- and four weeks post-intervention? (c) What are the students' perspectives on the early mathematics intervention?

The remainder of the paper is organized as follows. First, detailed background information regarding the T2BI and the CLPs, along with the procedures used to integrate CLPs and T2BI, are presented. The research design and data analysis are then discussed. Next, results obtained using the T2BI and CLPs are provided, revealing a functional relationship between the addition of CLPs to an existing intervention protocol and improved intervention response for four first-grade students with math difficulties. Finally, the implication of the result, the limitation, and recommendations for future research are discussed.

## Background Information

### Tier 2 Booster Intervention (T2BI)

The purpose of T2BI is to provide kindergarten, first-grade, and second-grade educators with the intervention lessons, teacher masters, student materials, and progress-monitoring tools needed to conduct the intervention with students identified as having MD. The T2BI mathematical content and skills are aligned with the Texas Essential Knowledge and Skills (TEKS) and include addition/subtraction combinations, word problem solving, number sequences, relationships of 10, and magnitude comparison. T2BI consist of 11 units, and each unit lasts for two weeks. In each 2-week period (unit), there are 16 lessons designed to teach the five skills across the eight-day unit and provide instruction at different levels to teach conceptual, procedural, and strategic knowledge.

The first page of each lesson identifies the concept and skill being taught, the name of the lesson, the objective, the instructional content (e.g., range of numbers), the materials, the vocabulary, and the instructional time (total time, instruction time, and time for independent practice). To help students develop conceptual understanding, content is represented in three ways in the lessons (i.e., concrete, pictorial, and abstract). A suggested script is provided to use

when implementing the lessons with students. Sidebars, or columns down the side of each lesson page, provide notes to teachers, student error-correction suggestions, and time boxes as instructional tips. Each day of instruction is estimated to take about 25 min, including warm-up, instruction, transition, and independent practice. The total time for each lesson, along with the time for instruction and independent practice, is listed at the top of the first page of each lesson. Similarly, timer icons throughout each lesson remind educators of the time allotted for particular sections.

### Cognitive Learning Principles (CLPs)

In previous reviews in the field of cognitive science (Dunlosky et al., 2013; Koedinger et al., 2013; Pashler et al., 2007), researchers have compared different types of CLPs and have found that one type might be more effective than the another depending on the instructional design, students' characteristics, learning conditions, materials and criterion tasks. In the following section, the CLPs for students' mathematics learning are described.

**Self-explanation** Self-explanation promotes sense-making and comprehension and has positive learning outcomes for students in various content areas (Booth et al., 2017; Koedinger et al., 2013). Research supports the effects of self-explanation on logical reasoning. The core component of self-explanation is asking students to explain some aspect of a problem during learning (Dunlosky et al., 2013). To assess the effects of self-explanation on logical reasoning, Berry (1983) compared the performance of three groups of students on the Wason card selection task. The problem-solving accuracy scores of two groups of students who self-explained while solving each problem and self-explained after solving all problems were more than students who were not prompted to engage in self-explanation (Berry, 1983). Also, self-explication is more effective in learning mathematics for students with MD than average-performing students (Kastens & Liben, 2007). For example, it facilitates learning simple addition problems for kindergarteners and mathematics equivalence problems for elementary-age students (Berry, 1983; Dunlosky et al., 2013).

**Interleaving practice** According to the interleaving practice principle, spreading out learning opportunities causes better long-term retention of information than providing multiple learning opportunities, one right after the other (i.e., massed practice; (Dunlosky et al., 2013). Researchers have found that students who received interleaving practice outperformed on the follow-up tests compared to their peers who received block practice (Taylor & Rohrer, 2010) asking students to solve mathematics problems on different content areas (addition, subtraction, place values) provides more

opportunities for students to detect their errors and refine their knowledge on different mathematics content areas (Li et al., 2012). In addition, interleaving practice can help students build a strong relationship between problem types and appropriate solution strategies (Rohrer et al., 2014).

**Error reflection** Learning from errors can be most effective if learners are encouraged to identify the features of the problem that caused them to make an error (i.e., find their mistake). This principle originated from cognitive dissonance theory, a state of tension or discomfort that arises whenever one holds two cognitions inconsistent with one another (Festinger, 1962). Research suggests that purposefully creating cognitive uncertainty can produce positive changes in students' thinking. Students better learn tasks that induce uncertainty and provoke them to resolve it, such as presenting information contradictory to their current knowledge (Overoye & Storm, 2015). Error reflection could be particularly relevant to mathematics because students' reflection on errors (either their errors or other learners' errors) leads to better understanding (Siegler & Chen, 2008). Several research studies have shown that studying and explaining errors can benefit mathematics learning for students with low and high prior knowledge (Barbieri & Booth, 2016; Heemsoth & Heinze, 2014). In addition, studying errors provides exposure to multiple perspectives rather than just one's perspective (Siegler & Chen, 2008).

**Worked examples** This principle suggests that asking learners to study examples of worked-out solutions to problems is more effective than asking them to solve all of the problems themselves (Sweller & Cooper, 1985). Studying worked examples can reduce learners' cognitive load by reducing the attentional and working memory demands needed to remember all of the problem-solving steps. Instead, they can focus their limited working memory capacity on understanding the reasoning behind the procedural steps taken in the example (Booth et al., 2015). Worked examples help students explain the steps in the example, connect new information to prior knowledge, and generate inferences to fill knowledge gaps (Mayer, 2014). Studying and explaining worked examples can be beneficial for novice learners (Booth et al., 2013). Studies showed that explaining worked examples improves conceptual understanding and sense-making in mathematics (Booth et al., 2015; Reed et al., 2013).

## Method

### Participants

Participants were recruited from a public elementary school in a district in central Texas in the fall. Most of the students

(55.2%) were economically disadvantaged based on free and reduced lunch status. Two first-grade classrooms (a total of 35 students) were screened using the school's universal screener and a norm-referenced measure to determine if they qualified for the study.

Multiple-gating procedures were utilized as cost-effective stepwise screening mechanisms to identify eligible participants (Loeber, 1990; Loeber et al., 1984). The first gate was using the results of the school-administered Texas Essential Knowledge and Skills (TEKS) aligned Pearson Education End of Year test (Envisions, 2014) from the school to identify students whose scores fell below the proficiency level (70% accuracy on the test) (DMAC Solutions, Education Service Center, 2018). Based on the Pearson Education End of Year test, 15 students fell below proficiency level and were nominated by their teachers to receive the interventions. The second gate was utilizing the standardized Test of Early Mathematics Ability, Third Edition (TEMA-3) (Ginsburg & Baroody, 2003) for students below the proficiency level on the Pearson Education End of Year test. After receiving consent forms from all 15 students, a mathematics interventionist administered TEMA-3 to identify students whose scores fell at or below the 30th percentile from the pool of students identified through the initial universal screening procedures. From the list of potential candidates, only students whose English was their first language (their parents are English speakers) were included in the study to have a homogeneous group. After excluding English Learners (ELs), only four participants were qualified to participate in the intervention study (i.e., Adam, Lucy, Sina, and Zara), all of whom received free or reduced lunch (participant names are pseudonyms). Students' demographics were collected from the school on one occasion in the fall semester. Table 1 provides the demographic data for the participants.

### Research Design

A multiple baseline design across participants (Kennedy, 2005) was implemented to assess the effects of the early mathematics intervention for students at risk for MD, utilizing progress monitoring measures. The basis of single-case research methodology relies upon repeated measurement of dependent variables before, during, and after introducing the independent variable to determine if a functional relation exists (i.e., a demonstration of experimental control) (Horner et al., 2005; Kennedy, 2005). Students who meet the criterion of needing mathematics intervention were assigned to the intervention sessions based on availability according to the general classroom schedule. Then, they were assigned to an order in which they received the intervention based on the baseline data. The four subtests of the Texas Early Mathematics Inventory–AIM Checks (TEMI-AC) were administered for the

**Table 1** Participant Demographic Information

		Adam	Lucy	Sina	Zara
Age (year/month)		6/9	7/3	7/2	7/2
Grade		1	1	1	1
Ethnicity		B	H	B	H
Gender		M	F	M	F
Universal Screener	Raw/PL	3/16	3/16	5/27	6/33
	Category	BA	BA	BA	BA
Screening: TEMA	Raw/PR	28/13	41/21	30/9	37/19
	Category	BA	BA	BA	BA

*Note.* F=Female; M=Male; H=Hispanic or Latino; B=Black or African American; PL=Proficiency Level; BA=Below Average; PR=Percentile Rank

baseline phase until a stable baseline was established (Horner et al., 2005). The intervention lasted six weeks for each student. During the intervention phase, The TEMI-AC was administered twice a week on Tuesday and Friday. The maintenance phase started after the conclusion of the last intervention session for each student. No further intervention sessions took place between the end of the intervention phase and the administration of maintenance measures. To assess maintenance, the TEMI-AC was administered to each of the participants during the typically scheduled intervention time two and four weeks after the final intervention sessions (Table 2).

**Measures**

Several measures were administered in this study. First, all students were given the TEMA-3 to determine whether they met the criteria for participation. Second, the TEMA-3 was also administered to determine the generalization effect. Finally, TEMI-AC was administered during the study's baseline, intervention, and maintenance phases to measure the functional relationship between the intervention and students' mathematics performance (Horner et al., 2005).

**TEMA-3**

The screening measure was TEMA-3 (Ginsburg & Baroody, 2003), a norm-referenced measure/diagnostic tool for determining mathematical strengths and weaknesses of students ages three through eight. The TEMA-3 was also utilized as a generalization measure in the post-intervention. The TEMA-3 consists of 72 items in the domains of informal and formal mathematics. Test results were reported as standard scores, percentile ranks, age, and grade equivalents. Internal consistency reliabilities were all above 0.92; immediate and delayed alternative form reliabilities were in the 0.80 s and 0.90 s (Ginsburg & Baroody, 2003). Students whose scores ranked at or below the 30th percentile were qualified to participate in the study.

**TEMI-AC**

The TEMI-AC contains four 2-min fluency measures assessing Magnitude Comparisons (circling, from two numbers shown, the number that is lower or circling both numbers if they are equal), Number Sequences (writing the number that is missing from a three-number sequence), Place Value (writing how many hundreds, tens, and ones are pictorially depicted), and Addition-Subtraction Combinations (solving basic addition and subtraction facts). The TEMI-ACs were aligned with the numerical skills and concepts taught in the intervention, which took approximately 10 min. The number and operation skills measured in the TEMI-ACs are essential for students to develop a foundation of number sense that is critical for later mathematics success (National Council of Teachers for Mathematics, 2014). The raw scores of the four measures were summed, yielding a total score that could be used to monitor student progress. The TEMI-AC has five alternate forms; the alternate-form reliability of the total score exceeds 0.80 across all forms. The measure was normed across 69 school districts in Texas in the fall. The percentiles range from 1 to 99; if a student scores at the 33rd percentile, 32 percent of the normative sample scored lower than the student, and 66 percent scored higher. The

**Table 2** Level and Trend Data for Participants

Participants	Level			Trend		
	Baseline	Intervention	Maintenance	Baseline	Intervention	Maintenance
Adam	13.57	53	59	-2.07	2.1	3.5
Lucy	61.72	87.08	90	0.63	-0.03	1.5
Sina	41.66	69.83	72	-0.83	2.85	-2
Zara	78.58	97.33	101.5	-0.13	0.31	0.5
Mean	48.88	76.81	80.63	-0.6	1.31	0.87
SD	27.96	19.51	18.85	1.14	1.38	2.29

*Note.* SD=Standard Deviation

cut-off point for the fall administration is the 25<sup>th</sup> percentile. If a student's score was below the 25<sup>th</sup> percentile, it is ranked as below average. The test validity for identification and progress-monitoring purposes was moderate to high ( $\alpha = 0.4$  to  $0.8$ ) (Bryant, 2008).

### Number Specific Knowledge

To assess number specific knowledge, three measures were administered before and after the intervention: an addition strategy task (the child was asked to solve problems on flashcards as quickly as possible without making too many mistakes), a number sets task (the child was asked to move across each line of the page from left to right without skipping any items, to circle any groups that could be put together to make the top number), and a number line estimation task (the child was asked to mark the line where the target number should lie) (Geary et al., 2018). In the number set sub-test, the participants were asked to move across each line of the page from left to right without skipping any items, to "circle any groups that can be put together to make the top number," and to "work as fast as you can without making many mistakes." There was a time limit of 60 s per page for target five and 90 s per page for target 9 (Geary et al., 2018). The overall frequency of hits and false alarms was calculated for each participant. The signal detection measure,  $d'$ , was calculated for each participant by subtracting the standardized number of false alarms from the standardized number of hits. The signal detection  $d'$  measure is the difference between standardized hits and standardized false alarms (Geary, Bailey, and Hoard, 2009). Geary et al. (2018) reported that the addition strategy, number set, and number line variables were highly correlated ( $r_s > 0.58$ ,  $p_s < 0.0001$ ). A confirmatory factor analysis with factor loadings constrained to equality confirmed that variables defined a single factor,  $\chi^2 = 0.33$ ,  $p = 0.84$ , goodness-of-fit index = 0.99 (Geary et al., 2018).

### Social Validity

To assess students' perspective about the early mathematics intervention, the first author developed a social validity survey that contained seven face scale questions (i.e., 3: happy face, 2: neutral face, and 1: a sad face.), two open-ended questions to express their thought toward the intervention, and one yes/ no question to see if they would volunteer to participate in this intervention again (see Appendix 1). The researcher verbally asked the social validity questions and recorded their answers. The mean score for the face scale questions was used to determine participants' perspectives toward the intervention.

### Fidelity of Implementation

A special education researcher observed the interventionist for four sessions (25% of intervention sessions) during the six weeks of intervention to assess implementation fidelity. The intervention fidelity was evaluated through four components, including tutors' ability to implement instructional; tutors' effectiveness in using explicit instruction (e.g., increasing scores); tutor's ability to promote students' verbalization by using questioning strategies; and the quality of the intervention (e.g., making the students feel valued and welcome; being responsible for the student's behaviors). Each component indicator was rated using a 3-point scale from poor to excellent, with one poor and three excellent (see Appendix 2).

### Procedures

We integrated four main effective CLPs in the design of 24 T2BI lessons, and we increased the dosage of T2BI by providing one to one 30–35 min of intervention sessions five days a week. The study took place in a quiet room at the school library every morning. The first author served as an interventionist and delivered all the intervention sessions. 25% of the intervention sessions were observed. Across all observations, the highest rating (3) was given on the level of interventionist competence. Overall, the intervention fidelity ranged from 85 to 100% throughout the intervention, with an average score of 92.5 (SD = 6.53).

### Instructional Design and Delivery

During the baseline phase, the participants attended their regular mathematics instruction class schedules. The TEMI-AC was administered to the participants weekly at approximately the same time as the school day when future intervention sessions were implemented. The intervention sessions start when a stable baseline (i.e., the data points are closer to the trend line) had been determined (Horner & Kratochwill, 2012; Kennedy, 2005) for the first student based on the total score of the TEMI-AC. Then, using multiple baseline procedures, intervention sessions began for each student in turn after achieving a stable baseline on the total score of the TEMI-AC.

### Instructional Component

The intervention consisted of 24 T2BI lessons that students received five days per week for six weeks. Each intervention session included a warm-up activity and a lesson. The CLPs that were embedded into the instruction include (a) self-explanation (i.e., encouraged students to share their thinking aloud about their solution approaches and their mathematical understanding); (b) interleaving practice (i.e., practicing the previously learned skills via flashcards); (c)

error reflection (i.e., encouraged students to identify their errors and think about what features of the problem make the specific step taken incorrectly); and (d) worked example (i.e., explaining worked out examples).

**Self-Explanation** During modeling and guided practice sections, the interventionist encouraged learners to generate an explanation for an explicitly stated fact, for example, ‘why does it make sense that...?’, ‘why is this true?’, ‘why?’, ‘why did you select that answer?’, ‘what steps did you take to get that answer?’, ‘why do you think it is true?’, ‘why do you think so?’. These questions help students connect their prior knowledge with the new information, organize information, and identify similarities and differences between related entities (Dunlosky et al., 2013).

**Interleaving Practice** At the beginning of each intervention, the interventionist presented a flashcard on the previously learned skills and asked the students to give a quick oral or written response (within five seconds). If students gave an incorrect answer to a flashcard, the interventionist put the card in a pile for extra practice. Spreading out learning opportunities leads to better long-term retention of information. In addition, the interleaving principle suggests that when practice problems are alternated, with a problem on one concept followed by a problem on another concept, students learn better than if problems are blocked or grouped by concept (Rohrer et al., 2012).

**Error Reflection** The interventionist provided corrective affirmative feedback and prompted students to identify their errors and think about what features of the problem make the specific step taken incorrectly. Studying errors provides exposure to multiple perspectives rather than just one’s perspective (Siegler & Chen, 2008). The interventionist encouraged students to think aloud about their solution approaches and mathematical understanding (Gersten et al., 2009).

**Worked Example** The interventionist showed multiple worked-out solutions to a problem and asked the student the following question: ‘suppose you are helping your teacher grade a math test. These are students’ responses; you need to decide which one is correct and incorrect and why’. Researchers have found that having learners study examples of worked-out solutions to problems is more effective for learning than solving all of the problems themselves (Sweller & Cooper, 1985).

### Instructional Content

The mathematical concepts and skills that were taught to each student in the intervention phase include (a) addition and subtraction combinations, (b) number sequences, (c) magnitude comparisons, and (d) relationships of 10. The content was aligned with the first-grade Texas Essential Knowledge and Skills (TEKS).

### Materials

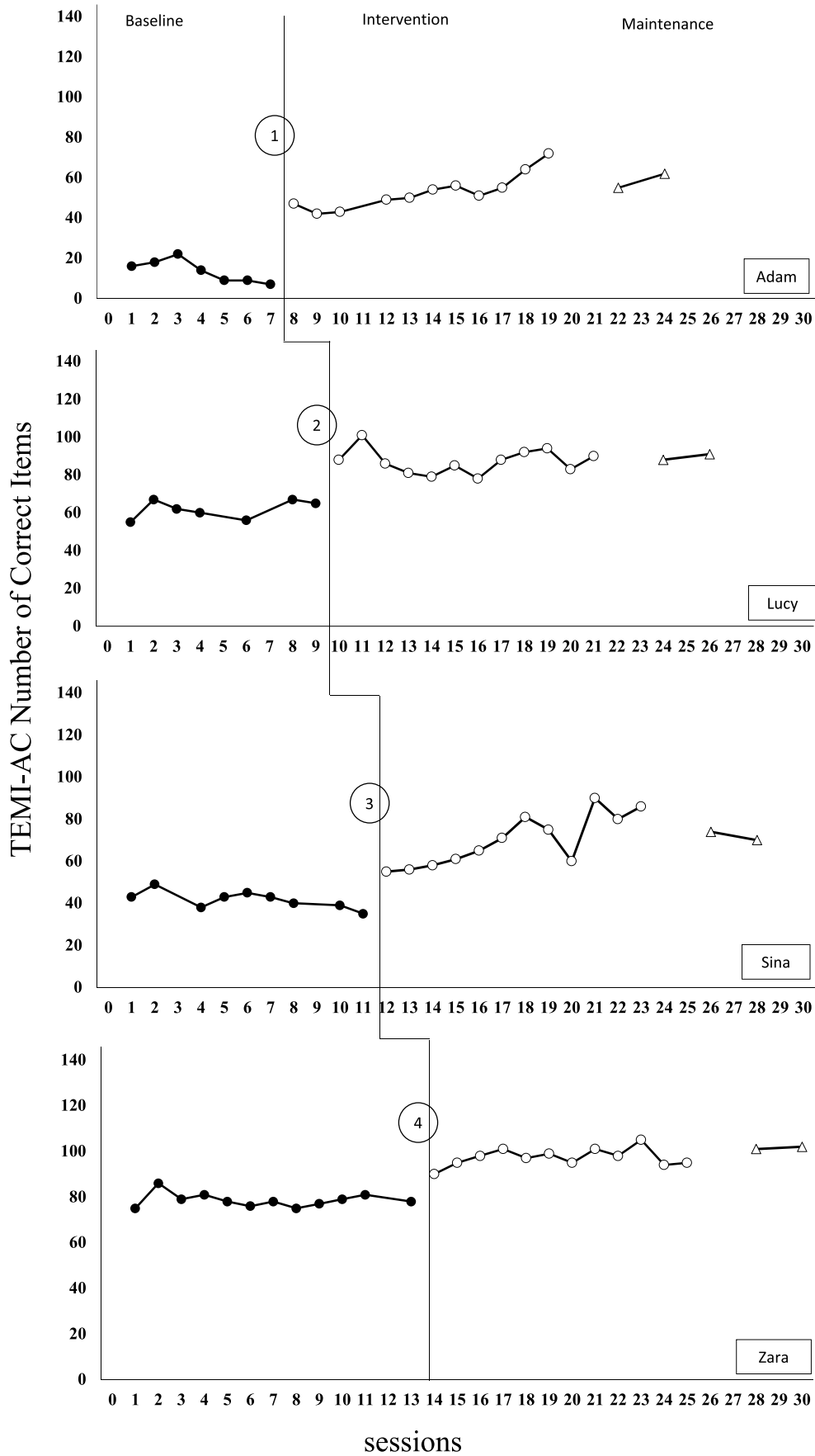
There were five types of materials: (a) hands-on materials that were used with lessons such as math manipulatives (e.g., connecting cubes, two-color counters, base ten blocks) and pictures that represented the concrete objects previously used; (b) templates or charts (e.g., hundred charts, ten-frames); (c) worksheets (i.e., modeled practice, guided practice, and independent practice); and (d) managing materials that the interventionist used to keep material organized such as a storage container with materials and lessons for each day. In addition, the interventionist used timers, wipe boards, as well as dry-erase markers. Every student received a booklet each day that contained all modeled practice, guided practice, and independent practice sheets that students needed in the lesson.

### Data Analysis

As recommended by What Works Clearinghouse (2014), six features were used to examine within- and between-phase data patterns to assess the effects of explicit strategic early mathematics intervention within single case design: (a) level, (b) trend, (c) variability, (d) immediacy of the effect, (e) overlap, and (f) consistency of data in similar phases (What Works Clearinghouse, 2014). The level of the data refers to the average of the data within each phase. The trend refers to the best-fit straight line placed over the data within each phase (Horner et al., 2005). The magnitude of the trend was calculated by the size or extent of the slope and qualitatively estimated as high, medium, and low (NAEP, 2019). The slope can be positive (upward) or negative (downward). Variability can also be defined as the degree to which the data points were dispersed relative to the best-fit straight line (Kratochwill et al., 2010). Finally, the immediacy of effect was determined by comparing the extent to which the level, trend, and variability of the last three data points in one phase were distinguishably different from the first three data points in the next (Kratochwill et al., 2010). Overlapping data refers to the percentage of data from one phase that overlaps with the data in the previous phase. Data consistency was identified by looking at data from all phases within the same condition (e.g., all baseline phases, all intervention phases) and identifying if there was consistency in the data patterns from phases with the same conditions (Kratochwill et al., 2010).

These six features were analyzed to determine if a causal relationship existed between the early explicit strategic mathematics intervention (i.e., independent variable) and the early numeracy knowledge and skills of students at risk for MD as exhibited through their performance (i.e., total score) on TEMI-AC (i.e., dependent variable).

We graphed and analyzed the data for the TEMI-AC score every week to determine baseline stability and intervention





◀**Fig. 1** Participants' TEMI-AC Total Scores by Session. The black dots indicate the baseline, the open dots represent the intervention phase, and the triangle symbols are for maintenance phase. The experimental control was established by the arrangement of conditions and manipulation of the independent variable across four different points in time

progress. We visually analyzed the level, trend, variability, immediacy of the effect, and consistency of data patterns on an ongoing basis as the study was executed. Additionally, statistical analysis was conducted using the non-overlap of all pairs (NAP) as a non-parametric effect size.

## Results

Research question one examined the effects of early mathematics intervention on the mathematics performance of first-grade students at risk for MD. The TEMI-AC was administered to assess the early mathematics concepts and skills of students at risk for MD. The authors used the total score of TEMI- AC to evaluate students' progress twice weekly.

## Visual Analysis

Figure 1 showed four demonstrations of predicted effect at four different points to assess experimental control in the study. First, the data from adjacent phases were compared, and then the data patterns from all phases in the study were integrated. To evaluate the effect across baseline and intervention phases (e.g., determine if the introduction of an intervention produced a predicted change in the early numeracy knowledge and skills), data from the second phase were compared initially with the data from the first phase and then with the "projected results" (e.g., the extension of the data pattern from the first phase into the second phase). In each case, data in the second phase were examined and compared (a) with the actual data from the first phase and (b) with the expected, or projected, data pattern (with confidence intervals) obtained by extending data from the first phase into the second phase (Fisher et al., 2011; Horner et al., 2005; Kratochwill et al., 2010). Visual data analysis involves simultaneously assessing the data's level, trend, and variability within adjacent phases. When data from two adjacent phases were compared, the rules of visual analysis also included assessment of immediacy of effect, the level of overlap, and the consistency of data patterns in similar phases (Parsonson & Baer, 1978).

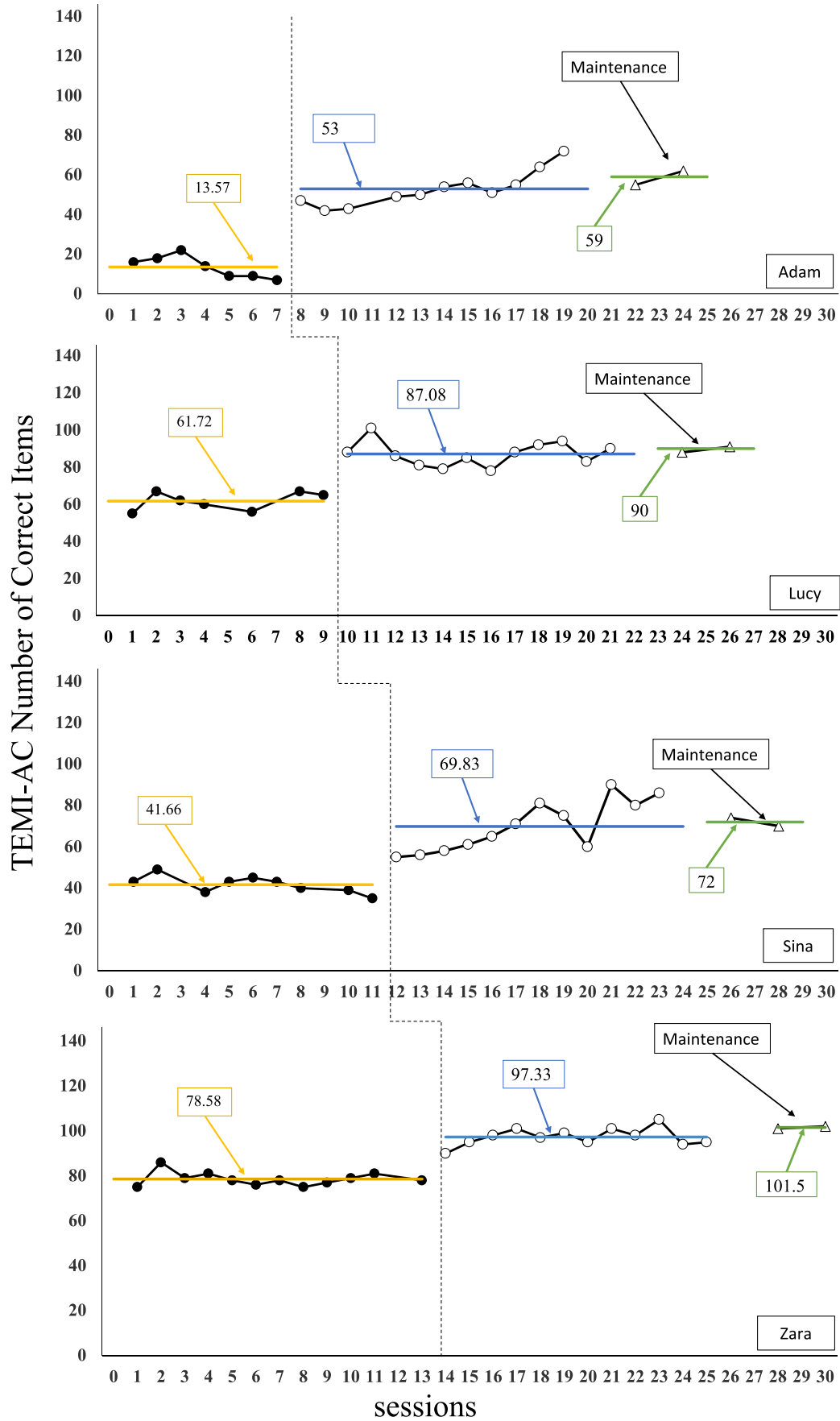
Active documentation of performance under baseline is the center feature of single-case design research. In this study, the baseline phase included between 7 to 13 data

points. The baseline data showed (a) the current pattern of responding and (b) a confident prediction of the pattern of future responding. The baseline data were collected individually in a quiet space in the school library, where later intervention sessions and progress monitoring happened. All the phases happened in the same place under the same condition to assure that only the independent variable was altered at the intervention point. All other baseline variables were held constant so that the independent variable was likely to be responsible for the change in the dependent variable (the early numeracy knowledge and skills). After introducing the intervention, all four participants improved their mathematical performance using the TEMI-AC total scores. The level data shows that the TEMI-AC total scores improved across the study and were maintained two and four weeks after the intervention was completed (see Fig. 2). Across all participants, the average level at baseline was 48.88 (SD = 27.96), which increased considerably during the intervention phase (M = 76.81, SD = 19.51). The trend analysis indicated that three participants showed a downward trend during the baseline (M = -0.6, SD = 1.14) and an upward trend during the intervention (M = 1.31, SD = 1.38) (see Fig. 3). The average variability across participants (i.e., the sum of all variables divided by 4) TEMI-AC total scores was 4.42 during baseline and 7.94 during the intervention (see Fig. 4).

The average immediacy of effect across all participants was 24.50 (SD = 9.54, range = 15– 35.66), meaning after the intervention, all participants showed high immediacy of effect from baseline to intervention phase (see Table 3). Furthermore, performance on TEMI-AC showed maintenance of scores at 2- and 4-weeks after the intervention across the participants. Based on the visual analysis findings, a causal relationship was demonstrated between implementing the explicit, systematic strategic early mathematics intervention and the mathematical performance of first-grade students at risk for MD.

## The Non-overlap of All Pairs (NAP)

NAP approach (Parker & Vannest, 2009) was computed as another method to assess the effectiveness of the intervention. Using the NAP approach, the authors examined the extent to which the TEMI-AC data in the baseline versus intervention and maintenance phases did overlap (see Table 3). NAP results were analyzed according to the following scale: 90–100% = large or highly effective, 70%–90% = moderately effective, and < 70% = small or questionable effectiveness. For all participants, the average possible pairs between baseline and intervention phases were 103.25 data points (range 77–144), and NAP was 100% which showed that from baseline to the intervention phase, data demonstrated a great improvement (Parker & Vannest, 2009). There was also no overlapped data point between phases across participants



◀**Fig. 2** Level for Students' TEMI-AC Total Score. The black dots indicate the baseline, the open dots represent the intervention phase, and the triangle symbols are for maintenance phase for each student. The number within each phase is the Level that refers to the average of the data within the phase

(NAP = 100%). The 100% NAP value demonstrates a large effect (Parker & Vannest, 2009) of explicit, systematic strategic early mathematics intervention across participants, verifying a causal relationship between the introduction of the intervention and changes in participants' mathematical performance on TEMI-AC (Horner et al., 2005) at four different time points. The explicit, systematic strategic early mathematics intervention on TEMI-AC scores can be interpreted as highly effective during the intervention and the maintenance phases compared to baseline. During the maintenance phase, NAP (100%) demonstrated large, long-lasting effects on TEMI-AC scores two and four weeks after the last intervention session.

### Maintenance

The TEMI-AC total scores increased across the study and were maintained two and four weeks after completing the intervention for each participant. Across all participants, the average level in baseline increased during the intervention phase and maintained during the maintenance phase ( $M = 80.63$ ,  $SD = 19.51$ ). The level of maintenance data was larger than both baseline and intervention phases for all participants (see Table 2). The level of maintenance data for Aaron, Lucy, Sina, and Zara was 59, 90, 72, and 101.5, respectively. Across the participants, the level of data increased at least 22 scores from the baseline phase to the maintenance phase. Results also demonstrated that students with a lower level in the baseline phase showed more improvement in the intervention phase and maintained the effect in the maintenance phase. Across all participants, the average level in baseline was 48.88 ( $SD = 27.96$ ), which improved considerably during the maintenance phase ( $M = 80.63$ ,  $SD = 19.51$ ).

### Generalization

To examine the generalization effect of the intervention, the TEMA-3 was administered as a distal measure in both pre and post-intervention. Table 4 shows the TEMA-3 result in the pre- and post-intervention. TEMA-3 was administered during the screening phase before the intervention, and students' mathematics abilities scores were below the average, ranging from 83 to 89 ( $M = 85$ ,  $SD = 4.40$ ). The average pre-test score across all students was 84.75, which is within the below-average category. The mean post-test score across participants was 91.25 ( $SD = 5.06$ ), falling within the average range. The results indicated that the intervention had a

significant effect on the overall mathematical performance of students at risk for MD (Table 5).

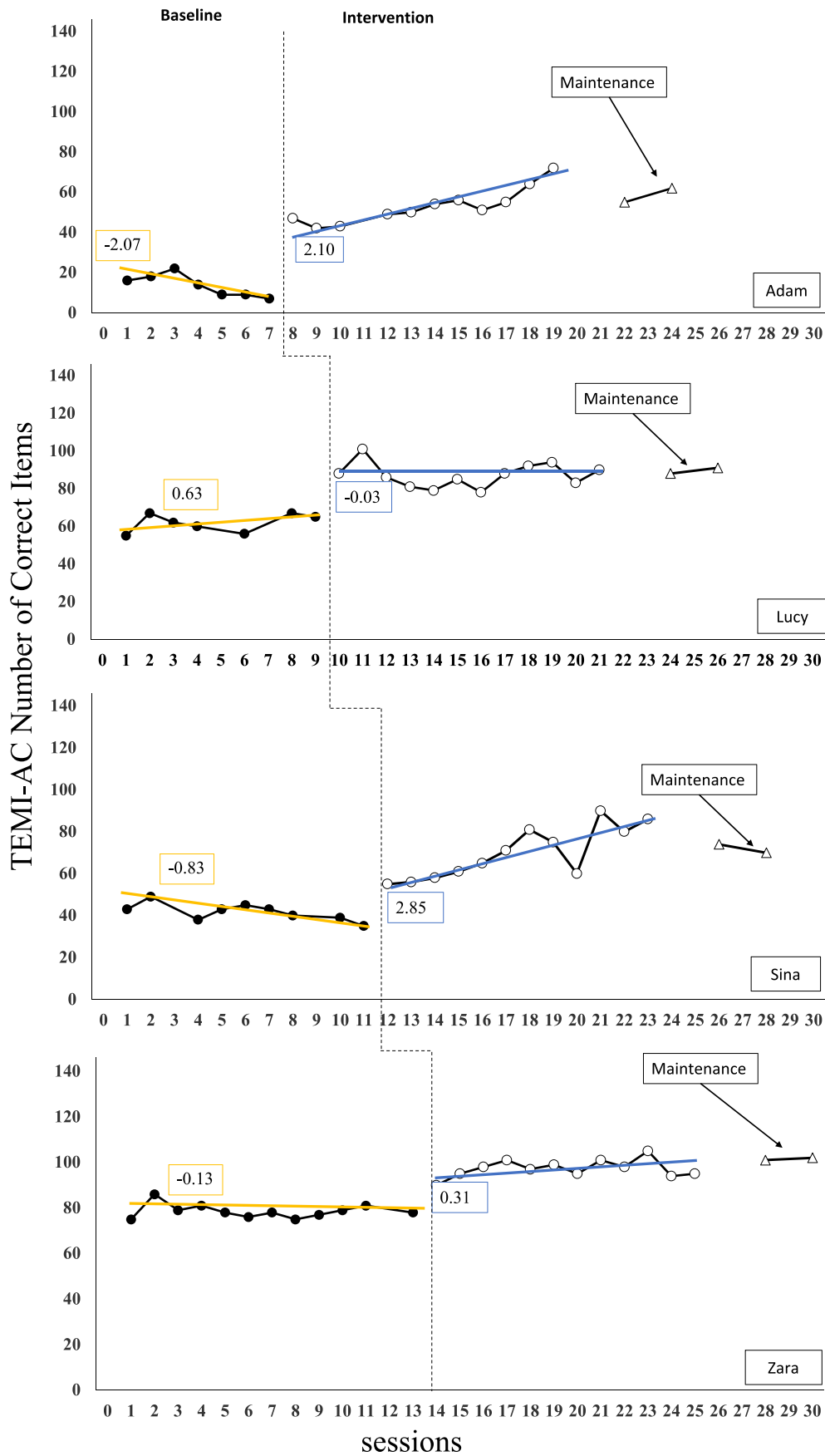
The results of the NSK measure demonstrated that participants' performance in the addition-strategy task improved after the intervention. Table 6 shows the percentage of using addition-strategy task strategies before and after the intervention. Table 7 shows that participants' performance in the number sets task improved in the post-test. The number-line estimation sub-test results indicated that there was a significant difference between the mean of differences in the pre-test ( $M = 18.87$ ,  $SD = 1.48$ ) and the post-test ( $M = 16.36$ ,  $SD = 0.59$ );  $t(6) = 3.48$ ,  $p < 0.05$ .

### Social validity

The social validity result showed that the average score of the first seven face scale questions was 2.75. All students expressed high levels of interest in participating in the program in the future. They also believed that the activities and lessons would help their classmates to do better in mathematics. All participants stated that learning mathematics is important and that the activities and lessons helped them perform better in their mathematics classes. Participants also believed that using different materials made mathematics skills easier to understand. The participants' favorite parts of the intervention were the flashcard activities, hands-on material, and independent practice worksheets.

### Discussion

The purpose of this study was to examine the effects of the integration of T2BI and CLPs on the performance of four first-grade students with MD. The intervention included the fundamental mathematical concepts and skills aligned with the first-grade Texas Essential Knowledge and Skills (TEKS), including addition/ subtraction combinations, number sequences, magnitude comparisons, and relationships. Previous studies demonstrated the positive effects of early mathematics intervention on the fundamental mathematics skills for students at risk for MD in early grades (Casey et al., 2008; Jordan et al., 2012; Klein et al., 2008; Sood & Jitendra, 2011). However, a significant subsample of students showed minimal response to the intervention or did not maintain the effects of the intervention. In this study, the T2BI and CLPs were effective in improving conceptual understanding and retaining mathematical knowledge. Participants in this study were identified as at risk for MD using a multiple-gating procedure, a cost-effective stepwise screening mechanism for identifying eligible participants (Loeber, 1990; Loeber et al., 1984). To examine the effect of the intervention, both visual analysis and proximal effect sizes (i.e., of visual data) were employed.



**Fig. 3** Trend for Students' TEMI-AC Total Score. The black dots indicate the baseline, the open dots represent the intervention phase, and the triangle symbols are for maintenance phase for each student. The best-fit line is presented over the data within each phase with a slope

The visual analysis results demonstrated a causal relationship between the systematic strategic mathematics intervention and participants' early mathematical knowledge and skills. All four participants demonstrated a lower level of mathematics performance during the baseline phase, but their level improved significantly from baseline to intervention. The participants' mathematical performance level increased after introducing the intervention, and all participants maintained the intervention effects two and four weeks following the intervention phase.

Additionally, all participants who qualified for this study come from low SES backgrounds. It is important to note that students from low SES backgrounds have been shown to have difficulties in early numeracy knowledge and skills compared to their peers from high socioeconomic backgrounds (Griffin et al., 1994; Jordan et al., 2012). The positive outcomes of this intervention were similar to previous research studies suggesting that early mathematics intervention positively affects students' mathematical abilities from economically disadvantaged backgrounds (Klein et al., 2008; Whyte & Bull, 2008). Also, The findings of this study were consistent with previous research; that is, explicit, systematic instruction is an effective instructional approach for improving students' mathematical skills with MD (Iseman & Naglieri, 2011; Van Luit & Naglieri, 1999).

The results of the maintenance effect of the systematic strategic early mathematics intervention demonstrated that all participants' TEMI-AC scores improvement maintained two and four weeks after the completion of the intervention phase. Therefore, the early mathematics intervention was effective for students at risk for MD, even after removing the intervention. The CLPs embedded in the intervention were possible variables that accounted for improving mathematics learning and maintaining the effects after the intervention. Dunlosky et al. (2013) reported that spreading out learning opportunities may result in better long-term retention of information (Dunlosky et al., 2013). Interleaving practice can improve retention by triggering elaborative retrieval processes because it involves searching for long-term memory that activates related information (Pashler et al., 2007).

The last research question examined the perspective of participants about the intervention. The findings of the social validity survey revealed that, on average, students had a positive perspective on the intervention. This finding suggests that the participant believes the intervention was helpful and socially valid. The results were aligned with previous studies showing positive perceptions of participants about the program's effectiveness and benefits (Calhoun et al., 2007; Jitendra et al., 2004).

## Educational implications

The intervention consisted of the explicit instructional components (e.g., warm-up, modeled practice, guided practice, independent practice, check for understanding, and error correction) and multiple CLPs for teaching mathematics. Teachers can use these instructional methods in the mathematics classroom to improve mathematical learning in early grades for students at risk for MD (Bryant et al., 2008; Fuchs et al., 2007). Also, all participants in this study were from low SES backgrounds. The results showed a significant effect of early mathematics intervention in this sample, aligned with previous research on early mathematics intervention (Klein et al., 2008; Siegler & Ramani, 2008). Thus, this study suggests that early mathematics intervention can be academically effective for students from low SES backgrounds.

## Limitations

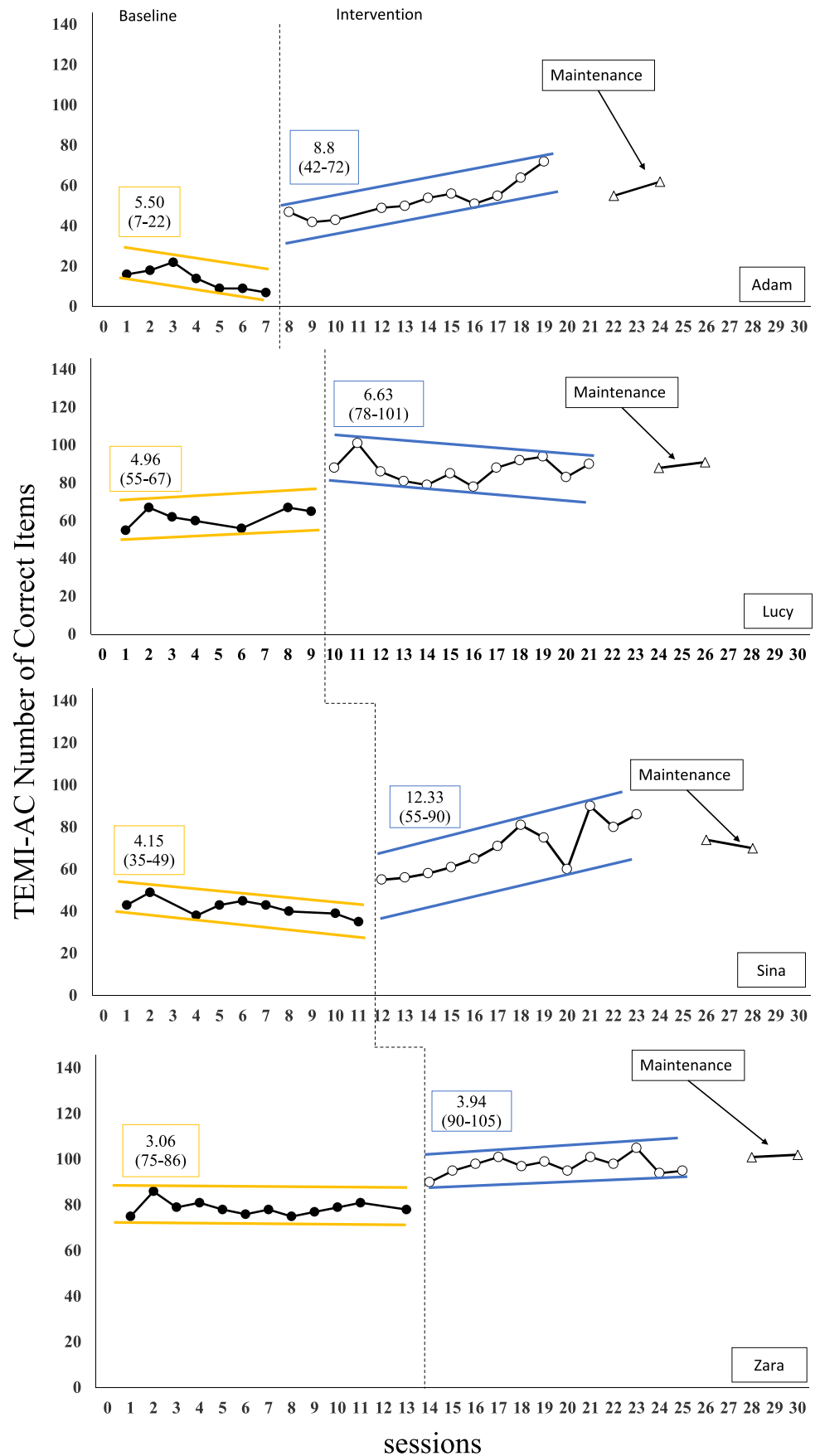
The first limitation is that the intervention included the components of explicit instruction (Gersten et al., 2009; Wang et al., 2016) and consisted of several CLPs for teaching mathematics (Dunlosky et al., 2013; Geary et al., 2017). Although the intervention effectively improved the mathematical knowledge and skills of students at risk for MD, it is not clear which intervention component is more effective for each mathematics concept. For example, the CLPs may vary in their effectiveness for simpler versus more complex mathematical content. Thus, CLPs embedded in the intervention design could contribute to the effects of the intervention. However, this claim was not directly examined in the current study, and further investigation is needed to purposefully test the effect of CLPs on students' mathematics achievement.

The second limitation is that the study was conducted with native English speakers who live in a large, urban city. All participants in the study were identified as at risk for learning disabilities and also were economically disadvantaged. Therefore, the results of this study cannot be generalized to English learners, students from different racial/ethnic backgrounds, or students who live in suburban or rural areas. Finally, we systematically evaluate the effects of the interventions in a one-on-one context using a single case design methodology. We recommend that future studies examine the effect of this intervention in a group format at a larger scale using randomized control trials (i.e., scaling up).

## Future research

First, the finding suggests that using multiple instructional components does not allow researchers to determine which principles could be more important than others. Although the CLPs were included in the design of mathematics interventions, these principles have not been explicitly manipulated to

**Fig. 4** Students' Variability and Immediacy of Effect. The black dots indicate the baseline, the open dots represent the intervention phase, and the triangle symbols are for maintenance phase for each student. The two parallel lines indicates the range of variabilities for each students within each phase



**Table 3** Variability, Immediacy of Effect, Overlap Data for Participants

Students	Variability (Range)		Immediacy of Effect	Overlap
	Baseline	Intervention		
Adam	5.50 (7–22)	8.84 (42–72)	35.66%	No
Lucy	4.96 (55–67)	6.63 (78–101)	29%	No
Sina	4.15 (35–49)	12.33 (55–90)	18.33%	No
Zara	3.06 (75–86)	3.94 (90–105)	15%	No
Mean (SD)	4.42 (1.06)	7.94 (3.55)	24.50% (9.54)	

Note. SD = Standard Deviation

**Table 4** TEMA-3 Total Standard Score

Participants	TEMA-3 Pre	TEMA-3 Post
Adam	83	85
Lucy	88	90
Sina	80	93
Zara	87	97
Mean (SD)	84.75 (4.31)	91.25 (5.06)

Note. TEMA-3 = Test of Early Mathematics Ability-3; SD = Standard Deviation

**Table 5** Participants' Pre-/Post-Intervention Total Scores in the Addition-Strategy Task

Participants	Addition-strategy task Pre	Addition-strategy task Post
Adam	29	39
Lucy	47	63
Sina	41	60
Zara	31	49
Mean (SD)	37 (8.49)	52.8 (11)

Note. SD = Standard Deviation

determine whether such principles effectively improve mathematical learning outcomes in young students at risk of MD. Perhaps struggling learners benefit the most from the rigorous application of such principles in instructional design. However, the researchers and educators have not purposefully tested these principles in applied intervention settings. So, more research needs to be done in this area.

**Table 6** Percentage of Accuracy and Sophistication of the Strategy

Accuracy and sophistication of the strategy	Adam		Lucy		Sina		Zara		Overall percentage	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Error in using the retrieval, fingers, or decomposition strategy	75	40	50	25	45	35	65	30	58.75	32.5
Error in using a counting strategy, whether finger or verbal counting	5	35	20	20	30	15	15	15	17.5	21.25
Correct use of the max or sum counting strategy	20	15	5	20	10	5	20	45	13.75	21.25
Correct use of the min counting strategy	0	10	10	10	10	25	0	5	5	12.50
Correct retrieval	0	0	15	25	5	20	0	5	5	12.50

Note. Pre = Before Intervention; Post = After Intervention

**Table 7** Participants' Pre-/Post-Intervention Scores in Number-Sets Task

Participants	Signal detection measure d' in pre-test	Signal detection measure d' in post-test
Adam	3	4.8
Lucy	5.4	6.4
Sina	3.4	7
Zara	4.6	5.8
Mean (SD)	4.1 (1.1)	6 (0.93)

Note. SD = Standard Deviation

Second, the results showed that students who started lower than others in baseline showed larger positive outcomes in intervention and maintenance phases. For example, Aaron, who had the lowest level in the baseline (53), showed the biggest change level (39.43) in the intervention phase. Based on the visual analysis of the data, we hypothesize that the severity of MD or prior knowledge could differentially impact students' responsiveness to the intervention. Further studies are required to examine this hypothesis. Also, we hypothesize that integrating a select set of validated CLPs into existing explicit, systematic intervention can accelerate mathematical learning and can promote the use of mature, efficient strategies in mathematics. We suggest that future research examine the effects of CLPs on the use of effective mathematics strategies.

Social validity surveys can determine whether those involved (e.g., teacher, parent, and student) believe that the

intervention is socially valid. The intervention implementation is more likely to be conducted as intended if those involved feel that the intervention is socially valid (Lane et al., 1999). For future research, we suggest using social validity surveys for parents, teachers, and students to better determine the social validity of the intervention.

Future research needs to examine the effect of teachers or school interventionists implementing this intervention to allow scaling up the intervention, which would involve

integrating the components into the routines of the teaching practices (Odom, 2008). Finally, although the participants' demographics represent part of the students' demographics identified as at risk, many EL students and students with different demographic characteristics may respond to the intervention differently. Future researchers may consider examining the intervention for students with different demographics and ELs.

## Appendix 1

Student Social Validity Scale.

SID: \_\_\_\_\_

Please read each statement and circle the comment that reflects your response.

**I really like Mathematics.**



**I think Mathematics is important.**



**The activities and lessons help me to do better in Mathematics class.**



**The activities we did helped me to better understand Mathematics concepts and skills.**



**The activities and lessons could help my classmates to do better in Mathematics class.**



**Using different materials made the skills easier to understand.**



**I feel as though I was able to finish many of the problems independently on the worksheets.**




**I would volunteer to participate in this program again YES or NO**



## Appendix 2

### Fidelity Check List.

 <div style="text-align: center;"> <h2 style="margin: 0;">Early Mathematics Intervention</h2> <h3 style="margin: 0;">Fidelity Checklist</h3> </div> <span style="float: right; background-color: #c00000; color: white; padding: 2px 5px; font-weight: bold;">2018</span>				
A. Session Info				
Interventionist		Time		
School		Lesson		
Date		Observer		
Type of fidelity check	<input type="checkbox"/> Live observation <input type="checkbox"/> Audio recording			
B. Interventionist Competence				
Elements	Score			Percentage
Shows good command and knowledge of subject matter	1	2	3	
Demonstrates breadth and depth of mastery				
Follows script, as necessary				
Has all needed materials prepared for the lesson				
Has few long pauses in execution of lesson				
Uses timer/recorder to monitor time spent on the session				
<b>sum</b>	1	2	3	
D. Explicit Instruction				
Elements	Score			Percentage
Provides sufficient cumulative review both before, during, and after instruction.	1	2	3	
Provides support to promote learning when concepts and skills are being first introduced to students.				
Watches and/or listens to student responses (of any form) to ensure that all students are mastering the skills that are being presented.				
Uses methods to promote active student response.				
Provides good models and uses think-alouds to make thinking visible and easier for students to follow.				
<b>sum</b>	0	0	0	0
D. Practice opportunities				
Elements	Score			Percentage
Provides varied opportunities for practice that are linked to the instruction	1	2	3	
Provides immediate, corrective and descriptive comments/suggestions for both correct and incorrect responses.				
<b>sum</b>	0	0	0	
E. Worked Example and self explanation				
Elements	Score			Percentage
Prompts students to explain the information to themselves.	1	2	3	
asks students to explain a combination of correct and incorrect examples				
Helps students to recognize and accept when they have chosen incorrect procedures				
Draws students' attention to the particular features in a problem that make the procedure inappropriate.				
For correct examples, promotes students to indicate "why" the step was right				
<b>sum</b>	0	0	0	0

**Data availability** The datasets generated during the current study are available from the corresponding author on reasonable request.

### Declarations

**Conflicts of interest** The authors have declared that no conflicts of interest exist.

**Ethical approval** The study was reviewed by the office of research support and compliance at the University of Texas at Austin. It was determined to be exempt from Institutional Review Board (IRB) review based on 45 CFR 46.101(b)(1, 2, 4) [IRB Exempt Determination for Protocol Number 2018–07-0134]. The authors certify that the study was performed following the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

**Informed consent** Written informed consent was obtained from the parents/ legal guardians for participation in the study and publication of the results.

### References

Al Otaiba, S., & Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention: A review of the literature. *Remedial and Special Education*. <https://doi.org/10.1177/07419325020230050501>

Barbieri, C., & Booth, J. L. (2016). Support for struggling students in algebra: Contributions of incorrect worked examples. *Learning and Individual Differences*, 48, 36–44. <https://doi.org/10.1016/j.lindif.2016.04.001>

- Barner, D., Alvarez, G., Sullivan, J., Brooks, N., Srinivasan, M., & Frank, M. C. (2016). Learning Mathematics in a Visuospatial Format: A Randomized, Controlled Trial of Mental Abacus Instruction. *Child Development*. <https://doi.org/10.1111/cdev.12515>
- Barnes, M. A., & Raghobar, K. P. (2017). Neurodevelopmental Disorders as Model Systems for Understanding Typical and Atypical Mathematical Development. In *Acquisition of Complex Arithmetic Skills and Higher-Order Mathematics Concepts* (pp. 67–97). Elsevier. <https://doi.org/10.1016/B978-0-12-805086-6.00004-7>
- Berch Geary, D. B., Ochsendorf, R., & Koepke, K. M. (2017). *Acquisition of complex arithmetic skills and higher-order mathematics concepts*. Elsevier Science.
- Berry, D. C. (1983). Metacognitive Experience and Transfer of Logical Reasoning. *The Quarterly Journal of Experimental Psychology Section A*, 35(1), 39–49. <https://doi.org/10.1080/14640748308402115>
- Booth, J. L., Lange, K. E., Koedinger, K. R., & Newton, K. J. (2013). Using example problems to improve student learning in algebra: Differentiating between correct and incorrect examples. *Learning and Instruction*, 25, 24–34. <https://doi.org/10.1016/j.learninstruc.2012.11.002>
- Booth, J. L., McGinn, K. M., Barbieri, C., Begolli, K. N., Chang, B., Miller-Cotto, D., Young, L. K., & Davenport, J. L. (2017). Evidence for Cognitive Science Principles that Impact Learning in Mathematics. In *Acquisition of Complex Arithmetic Skills and Higher-Order Mathematics Concepts* (pp. 297–325). Elsevier. <https://doi.org/10.1016/B978-0-12-805086-6.00013-8>
- Booth, J. L., Oyer, M. H., Paré-Blagoev, E. J., Elliot, A. J., Barbieri, C., Augustine, A., & Koedinger, K. R. (2015). Learning Algebra by Example in Real-World Classrooms. *Journal of Research on Educational Effectiveness*, 8(4), 530–551. <https://doi.org/10.1080/19345747.2015.1055636>
- Bryant, B. R. (2008). *Introduction to the Assessment Measures*. 25.
- Bryant, D. P., Bryant, B. R., Gersten, R. M., Scammacca, N., & Chavez, M. M. (2008). Mathematics intervention for first- and second-grade students with mathematics difficulties: The effects of tier 2 intervention delivered as booster lessons. *Remedial and Special Education*, 29(1), 20–32. <https://doi.org/10.1177/0741932507309712>
- Calhoon, B. M., Al Otaiba, S., Cihak, D., King, A., & Avalos, A. (2007). *Effects of a Peer-Mediated Program on Reading Skill Acquisition for Two-Way Bilingual First-Grade Classrooms* (Vol. 30, p. 184). <https://doi.org/10.2307/30035562>
- Casey, B. M., Andrews, N., Schindler, H., Kersh, J. E., Samper, A., Copley, J., Casey, B. M., Andrews, N., Schindler, H., Kersh, J. E., Casey, B. M., Andrews, N., Schindler, H., Kersh, J. E., Samper, A., & Copley, J. (2008). *The Development of Spatial Skills Through Interventions Involving Block Building Activities*. *The Development of Spatial Skills Through Interventions Involving Block Building Activities*. 0008. <https://doi.org/10.1080/0737000802177177>
- Clarke, B., Doabler, C. T., Cary, M. S., Kosty, D., Baker, S., Fien, H., & Smolkowski, K. (2014). Preliminary evaluation of a Tier 2 mathematics intervention for first-grade students: Using a theory of change to guide formative evaluation activities. *School Psychology Review*
- Clarke, B., Doabler, C. T., Kosty, D., Nelson, E. K., & Smolkowski, K. (2017). *Testing the Efficacy of a Kindergarten Mathematics Intervention by Small Group Size*, 3(2), 1–16. <https://doi.org/10.1177/2332858417706899>
- Compton, D. L., Gilbert, J. K., Jenkins, J. R., Fuchs, D., Fuchs, L. S., Cho, E., Barquero, L. A., & Bouton, B. (2012). Accelerating Chronically Unresponsive Children to Tier 3 Instruction: What Level of Data Is Necessary to Ensure Selection Accuracy? *Journal of Learning Disabilities*. <https://doi.org/10.1177/0022219412442151>
- DMAC Solutions, Education Service Center. (2018).
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest*, 14(1), 4–58. <https://doi.org/10.1177/1529100612453266>
- enVisionMATH Common Core (2014).
- Festinger, L. (1962). *A theory of cognitive dissonance* (Vol. 2). Stanford university press.
- Fisher, Piazza, & C. C., & Roane, H. S. . (2011). *Handbook of applied behavior analysis*. Guilford Press.
- Fuchs, D., Fuchs, L., & Vaughn, S. (2014). What Is Intensive Instruction and Why Is It Important? *TEACHING Exceptional Children*, 46(4), 13–18. <https://doi.org/10.1177/0040059914522966>
- Fuchs, L. S., Fuchs, D., & Hollenbeck, K. N. (2007). Extending Responsiveness to Intervention to Mathematics at First and Third Grades. *Learning Disabilities Research & Practice*. <https://doi.org/10.1111/j.1540-5826.2007.00227.x>
- Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Schatschneider, C., Hamlett, C. L., et al. (2013). Effects of First-Grade Number Knowledge Tutoring With Contrasting Forms of Practice. *Journal of educational psychology*, 105(1), 58–77. <https://doi.org/10.1037/a0030127>
- Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology*, 47(6), 1539–1552. <https://doi.org/10.1037/a0025510>
- Geary, D. C., Bailey, D. H., & Hoard, M. K. (2009). Predicting Mathematical Achievement and Mathematical Learning Disability with a Simple Screening Tool: The Number Sets Test. *Journal of Psychoeducational Assessment*, 27(3), 265–279. <https://doi.org/10.1177/0734282908330592>
- Geary, D. C., vanMarle, K., Chu, F. W., Rouders, J., Hoard, M. K., & Nugent, L. (2018). Early Conceptual Understanding of Cardinality Predicts Superior School-Entry Number-System Knowledge. *Psychological Science*. <https://doi.org/10.1177/0956797617729817>
- Gersten, R. M., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics Instruction for Students With Learning Disabilities: A Meta-Analysis of Instructional Components. *Review of Educational Research*, 79(3), 1202–1242. <https://doi.org/10.3102/0034654309334431>
- Gersten, R. M., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities*, 38(4), 293–304. <https://doi.org/10.1177/00222194050380040301>
- Gersten, R. M., Rolfhus, E., Clarke, B., Decker, L. E., Wilkins, C., & Dimino, J. (2015). Intervention for First Graders With Limited Number Knowledge: Large-Scale Replication of a Randomized Controlled Trial. *American Educational Research Journal*. <https://doi.org/10.3102/0002831214565787>
- Ginsburg, H., & Baroody, A. J. (2003). *The test of early mathematics ability*. Pro-ed.
- Griffin, S. A., Case, R., & Siegler, R. S. (1994). Right start: Providing the central conceptual prerequisites for first formal learning of arithmetic to students at risk for school failure. In *Classroom Lessons: Integrating Cognitive Theory and Classroom Practice*.
- Heemsoth, T., & Heinze, A. (2014). The impact of incorrect examples on learning fractions: A field experiment with 6th grade students. *Instructional Science*, 42(4), 639–657. <https://doi.org/10.1007/s11251-013-9302-5>
- Horner, R. H., Carr, E. G., Halle, J., Mcgee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*. <https://doi.org/10.1177/001440290507100203>
- Horner, R. H., & Kratochwill, T. R. (2012). Synthesizing Single-Case Research to Identify Evidence-Based Practices: Some Brief Reflections. *Journal of Behavioral Education*, 21(3), 266–272.

- Iseman, J. S., & Naglieri, J. A. (2011). A cognitive strategy instruction to improve math calculation for children with ADHD and LD: A randomized controlled study. *Journal of Learning Disabilities*. <https://doi.org/10.1177/0022219410391190>
- Jitendra, A., Edwards, L., Starosta, K., Sacks, G., Jacobson, L., & Choutka, C. (2004). *Early Reading Instruction for Children with Reading Difficulties: Meeting the Needs of Diverse Learners* (Vol. 37, p. 439). <https://doi.org/10.1177/00222194040370050501>
- Jordan, N. C., Glutting, J., Dyson, N., Hassinger-das, B., & Irwin, C. (2012). *Building Kindergartners' Number Sense: A Randomized Controlled Study*, 104(3), 647–660. <https://doi.org/10.1037/a0029018>
- Kastens, K. A., & Liben, L. S. (2007). Eliciting self-explanations improves children's performance on a field-based map skills task. *Cognition and Instruction*. <https://doi.org/10.1080/073700007019336702>
- Kearns, D. M., & Fuchs, D. (2013). Does Cognitively Focused Instruction Improve the Academic Performance of Low-Achieving Students? *Exceptional Children*, 79(3), 263–290. <https://doi.org/10.1177/001440291307900200>
- Kennedy, C. H. (2005). *Single-case designs for educational research*. Pearson/A & B.
- Klein, A., Starkey, P., Clements, D., Sarama, J., & Iyer, R. (2008). Effects of a pre-kindergarten mathematics intervention: A randomized experiment. *Journal of Research on Educational Effectiveness*, 1(3), 155–178. <https://doi.org/10.1080/19345740802114533>
- Koedinger, K. R., Booth, J. L., & Klahr, D. (2013). Instructional complexity and the science to constrain it. *Science*. <https://doi.org/10.1126/science.1238056>
- Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). What works Clearinghouse: Single-Case Design Technical Documentation. *What Works Clearing House*.
- Lane, K. L., Umbreit, J., & Beebe-Frankenberger, M. E. (1999). Functional Assessment Research on Students with or at Risk for EBD: 1990 to the Present. *Journal of Positive Behavior Interventions*, 1(2), 101–111. <https://doi.org/10.1177/109830079900100204>
- Li, N., Cohen, W. W., & Koedinger, K. R. (2012). *Problem Order Implications for Learning Transfer BT - Intelligent Tutoring Systems (S. A. Cerri, W. J. Clancey, G. Papadourakis, & K. Panourgia, Eds.; pp. 185–194)*. Springer Berlin Heidelberg.
- Loeber, R. (1990). Development and risk factors of juvenile antisocial behavior and delinquency. *Clinical Psychology Review*. [https://doi.org/10.1016/0272-7358\(90\)90105-J](https://doi.org/10.1016/0272-7358(90)90105-J)
- Loeber, R., Dishion, T. J., & Patterson, G. R. (1984). Multiple gating: A multistage assessment procedure for identifying youths at risk for delinquency. *Journal of Research in Crime and Delinquency*. <https://doi.org/10.1177/0022427884021001002>
- Mayer, R. (2014). *The Cambridge Handbook of Multimedia Learning (2nd ed.)*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369>
- Morsanyi, K., Devine, A., Nobes, A., & Szudouble acutecs, D. (2013). The link between logic, mathematics and imagination: Evidence from children with developmental dyscalculia and mathematically gifted children. *Developmental Science*. <https://doi.org/10.1111/desc.12048>
- National Center for Education Statistics. (2019). National Assessment of Educational Progress. *Encyclopedia of Education Economics & Finance*. <https://doi.org/10.4135/9781483346595.n176>
- National Council of Teachers of Mathematics. (2014). *Principles to actions*. National Council of Teachers of Mathematics.
- Odom, S. L. (2008). The Tie That Binds: Evidence-Based Practice, Implementation Science, and Outcomes for Children. *Topics in Early Childhood Special Education*, 29(1), 53–61. <https://doi.org/10.1177/0271121408329171>
- Overoye, A. L., & Storm, B. C. (2015). Harnessing the power of uncertainty to enhance learning. *Translational Issues in Psychological Science*, 1(2), 140–148. <https://doi.org/10.1037/tps0000022>
- Parker, R. I., & Vannest, K. (2009). An Improved Effect Size for Single-Case Research: Nonoverlap of All Pairs. *Behavior Therapy*. <https://doi.org/10.1016/j.beth.2008.10.006>
- Parsonson, B. S., & Baer, D. M. (1978). Training generalized improvisation of tools by preschool children1. *Journal of Applied Behavior Analysis*. <https://doi.org/10.1901/jaba.1978.11-363>
- Pashler, H., Rohrer, D., Cepeda, N. J., & Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin & Review*, 14(2), 187–193.
- Reed, S. K., Corbett, A., Hoffman, B., Wagner, A., & MacLaren, B. (2013). Effect of worked examples and Cognitive Tutor training on constructing equations. *Instructional Science*. <https://doi.org/10.1007/s11251-012-9205-x>
- Roediger, H. L., & Pyc, M. A. (2012). Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice. *Journal of Applied Research in Memory and Cognition*, 1(4), 242–248. <https://doi.org/10.1016/j.jarmac.2012.09.002>
- Rohrer, D., Dedrick, R. F., & Burgess, K. (2014). The benefit of interleaved mathematics practice is not limited to superficially similar kinds of problems. *Psychonomic Bulletin & Review*, 21(5), 1323–1330. <https://doi.org/10.3758/s13423-014-0588-3>
- Shrager, J., & Siegler, R. S. (1998). SCADS: A Model of Children's Strategy Choices and Strategy Discoveries. *Psychological Science*, 9(5), 405–410. <https://doi.org/10.1111/1467-9280.00076>
- Siegler, R. S., & Chen, Z. (2008). Differentiation and integration: Guiding principles for analyzing cognitive change. *Developmental Science*. <https://doi.org/10.1111/j.1467-7687.2008.00689.x>
- Siegler, R. S., & Ramani, G. B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental Science*, 11(5), 655–661. <https://doi.org/10.1111/j.1467-7687.2008.00714.x>
- Sood, S., & Jitendra, A. K. (2011). An Exploratory Study of a Number Sense Program to Develop Kindergarten Students' Number Proficiency. *Journal of Learning Disabilities*, 46(4), 328–346. <https://doi.org/10.1177/0022219411422380>
- Strand Cary, M. G., Clarke, B., Doabler, C. T., Smolkowski, K., Fien, H., & Baker, S. K. (2017). A Practitioner Implementation of a Tier 2 First-Grade Mathematics Intervention. *Learning Disability Quarterly*, 40(4), 211–224.
- Sweller, J., & Cooper, G. A. (1985). The Use of Worked Examples as a Substitute for Problem Solving in Learning Algebra. *Cognition and Instruction*, 2(1), 59–89. [https://doi.org/10.1207/s1532690xc0201\\_3](https://doi.org/10.1207/s1532690xc0201_3)
- Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology*, 24(6), 837–848. <https://doi.org/10.1002/acp.1598>
- U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) (2019). Mathematics Assessment.
- Van Luit, J. E. H., & Naglieri, J. A. (1999). Effectiveness of the MASTER program for teaching special children multiplication and division. *Journal of Learning Disabilities*. <https://doi.org/10.1177/002221949903200201>
- Vaughn, S., Wanzek, J., Murray, C. S., Scammacca, N., Linan-Thompson, S., & Woodruff, A. L. (2009). Response to early reading intervention: Examining higher and lower responders. *Exceptional Children*. <https://doi.org/10.1177/001440290907500203>
- Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Chen, R., Pratt, A., & Denckla, M. B. (1996). Cognitive Profiles of Difficult-to-Remediate and Readily Remediated Poor Readers: Early Intervention as a Vehicle for Distinguishing between Cognitive and Experiential Deficits as Basic Causes of Specific Reading

- Disability. *Journal of Educational Psychology*. <https://doi.org/10.1037/0022-0663.88.4.601>
- Wang, A. H., Firmender, J. M., Power, J. R., & Byrnes, J. P. (2016). Understanding the Program Effectiveness of Early Mathematics Interventions for Prekindergarten and Kindergarten Environments: A Meta-Analytic Review. *Early Education and Development*, 27(5), 692–713. <https://doi.org/10.1080/10409289.2016.1116343>
- What Works Clearinghouse. (2014). *What Works Clearinghouse Procedures and Standards Handbook*. 1–86. <https://doi.org/10.1037/e578392011-004>
- Whyte, J. C., & Bull, R. (2008). Number games, magnitude representation, and basic number skills in preschoolers. *Developmental Psychology*, 44(2), 588–596. <https://doi.org/10.1037/0012-1649.44.2.588>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.