aperio

RiSE
RESEARCH IN SPECIAL EDUCATION

ALETHEIA
S O C I E T Y

EMPIRICAL ARTICLE

Exploring the Need for Response-Guided Changes to Modified Schema-Based Instruction for Students with Intellectual Disability

Deidre Gilley*, Educational Psychology, Baylor, deidre_gilley@baylor.edu

Jenny Root, Florida State University, jrroot@fsu.edu

Alicia Saunders, School of Teacher Education, Florida State University, afsaunders@fsu.edu

Sarah Cox, Michigan State University, skcox@msu.edu

Danielle Morsching, Georgia Southern University, Dmorsching@georgiasouthern.edu

*Corresponding author.

ABSTRACT

This article reports findings from a feasibility study of three units of Math Scene Investigators, a curriculum that uses Modified Schema-Based Instruction (MSBI) to teach multiplicative word problem solving. The intent was to better understand *for whom* and *under what conditions* MSBI is effective, by examining how four high school students with intellectual disability (ID) responded to the intervention *as planned* and the need for and impact of response-guided changes. Results of the non-concurrent multiple baseline design found a functional relation

between MSBI and measured word problem solving behaviors for both ratio and proportion word problems. Two participants required response-guided changes during the proportion unit. We discuss how the results align with our anticipated data patterns and how post-hoc exploratory analysis informed our interpretation of participants' need for and response to these changes.

Keywords: mathematics; evidence-based practice; autism spectrum disorder; intellectual disability; single-case design

All students should learn to solve word problems because they provide an opportunity to practice applying mathematical knowledge and skills to everyday situations. Schema instruction is a well-researched instructional approach for teaching word problem solving that has over two decades of evidence showing positive impacts for students with learning disabilities and mathematics difficulties (Lein et al., 2020). Drawing on schema-induction theory (Gick & Holyoak, 1983), schema instruction in mathematics explicitly teaches students to identify mathematical structures (schemas) within a word problem and make connections between novel and familiar word problems (Fuchs et al., 2004). According to Powell and Fuchs (2018), schema instruction involves teaching students: (a) how to identify and classify word problems based on underlying mathematical structure (i.e., schema or problem type), (b) an efficient solution strategy for each schema, and (c) important vocabulary and language of word problems.

Browder and colleagues sought to investigate whether schema instruction could be effective for students who access the general curriculum through their state's alternate achievement standards linked to grade-level content standards but are reduced in complexity, breadth, and depth. Despite the commonality of having a "significant cognitive disability" in order to qualify to take their state's alternate assessment aligned with alternate achievement standards, variability is a defining feature of this group of learners. For example, while some students may use speech to communicate their reasoning, alternate means of action and expression such as augmentative and alternative communication (AAC) is necessary for those with complex communication needs. Browder and colleagues concluded that the core components of schema instruction alone may not sufficiently address the barriers to problem solving these students are likely to experience but could be effective if combined with established evidence-based practices.

INTENSIFICATION OF SCHEMA INSTRUCTION TO MODIFIED SCHEMA-BASED INSTRUCTION

The heterogeneous learning profiles of students determined to have a "significant cognitive disability" were the catalyst behind Browder and colleagues' work on intensifying schema instruction through a four-year research and development grant funded by the National Center for Special Education Research. Their research teams conducted eight single-case studies between 2013 and 2017 with elementary and middle school students with autism spectrum disorder (ASD) and intellectual disability (ID). They synthesized findings into a proposed model for modified schema-based instruction (MSBI; Spooner et al., 2017) that intensified the three core components of schema instruction to flexibly meet the heterogeneous needs of these students in four areas: (a) accessing the problem, (b) demonstrating conceptual understanding of the mathematical relationship depicted in the problem, (c) procedurally solving the problem, and (d) meaningful generalization. Root et al. (2020) expanded on the conceptual model by demonstrating how intentional application of the Universal Design for Learning framework (UDL Guidelines 2.0; CAST, 2018) enhanced the flexibility of MSBI to proactively plan for learner variability by considering barriers in the learning environment.

Accessing the Problem

The first aspect of MSBI that distinguishes it from other forms of schema instruction is attention to barriers in accessing the problem, including reading level, problem structure, quantities/content, and vocabulary (Spooner et al., 2017). The flexibility of MSBI is critical here, as the individual supports needed exist on a continuum and likely differ within and between students based on the task and their phase of learning (Root et al., 2022).

Conceptual Understanding

As a specific type of schema instruction, MSBI teach students to use schematic diagrams (i.e., graphic organizers) to model their conceptual understanding of each schema. In contrast to other forms of schema instruction wherein students copy or draw schematic diagrams, MSBI provides physical or virtual schematic diagrams with tailored visual supports such as color-coding and icons paired with text as antecedent cues. A universally designed schematic diagram enables students to model the relationship represented in the problem using concrete, representational, and abstract methods regardless of their communication, language, or fine motor skills.

Procedural Supports

In schema instruction, students often learn a mnemonic to remember an attack strategy consisting of logically sequenced steps for understanding and solving word problems, such as "FOPS" (F̲ind the problem, O̲rganize information using a diagram, P̲lan to solve the problem, and S̲olve the problem; Jitendra & Star, 2012). The level of literacy skills necessary to successfully use a mnemonic-based strategy is a barrier for many students. Relatedly, multiple interdependent decisions are encompassed within a step like "Solve the Problem." In MSBI, a student-friendly task analysis presents the attack strategy as a sequence of distinct, discrete behaviors in the form of a checklist that may include visual supports as needed.

Meaningful Generalization

The guidelines for word problems put forward by Spooner et al. (2017) emphasize making explicit connections to real-life activities and routines by providing multiple examples of each schema within personally relevant contexts. MSBI mediates response generalization by teaching a "rule" for each schema, a metacognitive strategy consisting of a verbal phrase (e.g., "Small group and small group combine into big group.") with corresponding hand gestures (e.g., bringing two fists together to represent a part-part-whole relationship).

Response-Guided Changes to MSBI

Given the individualized and flexible nature of MSBI and heterogeneity of this population, researchers have primarily used single-case experimental designs (SCED) to evaluate its impact on word problem solving (Root et al., 2021). One unique feature of SCED that is not shared by group experimental designs is the ability to make changes to specific aspects of an intervention based on participant data (Tincani & Travers, 2022). This flexibility has allowed researchers to maintain experimental control while simultaneously tailoring MSBI using the UDL framework and Spooner et al.'s (2017) conceptual framework.

Published MSBI studies have consistently reported response-guided changes based on categorization of data as indicating a skill or performance deficit (Cox et al., 2024). Changes have been made to stimuli (e.g., materials, task analysis format), behavioral supports (e.g., self-management strategies, individualized reinforcement systems), and instructional strategies (e.g., massed trials, system of least prompts). For example, Root et al. (2020) determined the variable data of one middle school student with ID was not due to a skill deficit (i.e., he knew what was expected) but rather a performance deficit stemming from a reluctance to talk aloud. Authors drew on the UDL framework to remove this barrier within the task by teaching an alternative means of action and expression (pointing).

A second example is seen in Cox et al. (2024). They attributed stagnant performance of two high school participants with ASD/ID to a skill deficit, as they were both making consistent errors reflecting procedural misunderstandings in calculating percent of change. In response, interventionists increased intensity of instruction by providing step-level feedback using a system of least prompts rather than waiting to deliver all feedback after the entire problem was solved independently. In contrast, Root et al. (2020) categorized the stagnant performance of a middle school student with ASD/ID as a performance deficit, indicating the intervention as planned was not sufficiently engaging because the participant was successful during guided practice but rushed through independent practice problems. The researchers and teacher collaboratively implemented an alternate contingent reinforcement system to increase motivation.

Two literature reviews on MSBI indicate that these examples are not isolated cases. Root and colleagues (2021) discussed this as an area for future research to understand "under what conditions" MSBI is effective. Relatedly, Clausen et al. (2021) concluded more evidence is needed to demonstrate how to support students who have limited literacy or complex communication needs. Understanding when, why, and how response-guided changes should be made is particularly relevant to MSBI, as it can provide a framework for using the principles of UDL to identify barriers and implement aligned supports.

Current Study

This manuscript reports findings from a feasibility test of three units of Math Scene Investigators, a word problem solving curriculum that uses MSBI to teach multiplicative schemas to secondary students whose mathematical instruction is based on alternate achievement standards. The primary purpose was to examine the effectiveness of MSBI *as planned* by asking "Is there a functional relation between MSBI and word problem solving behaviors?" The secondary purpose was to explore the need for and impact of response-guided changes to MSBI.

METHOD

A non-concurrent multiple baseline (NCMB) across participants design evaluated the effect of MSBI on word problem solving behaviors. The design, implementation, analysis, and reporting of findings followed Ledford et al.'s (2023) recommended guidelines for SCED.

Participants and Setting

The study took place in a large suburban public high school located in the southeastern United States. Researchers asked one special education teacher who taught a mathematics course aligned to the state's alternate achievement standards to send consents home with all of his students who were able to communicate their basic wants and needs in English either vocally or with AAC. Although researchers did not have access to educational records, all of his students received academic instruction aligned to alternate achievement standards, indicating they met the state's criteria for a "significant cognitive disability." In total, four students returned consents.

Researchers screened all four consented students using a researcher-created prescreening tool to assess the following skills: (a) receptive and expressive identification of double-digit numbers, (b) transferring numbers to a calculator, (c) using a calculator to solve single- and double-digit equations with whole numbers, and (d) solving ratio/proportion word problems. Table 1 details demographics of the four students with ID who were invited to participate in this study based on their performance on the prescreening tool. All invited participants demonstrated a basic level of prerequisite mathematics skills (a–c) during the screening as well as a need for the intervention (d). Three participants previously participated in studies with the research team that targeted different schemas between one and three years prior to the

onset of the current study (Phillip and Esteban, multiplicative comparison; Jayla, percent of change). We did not expect generalization to ratio and proportion word problems given performance on the prescreening (d).

	SEX	GRADE	RACE/ETHNICITY	PRIMARY ELIGIBILITY CATEGORY	SECONDARY ELIGIBILITY CATEGORY
Phillip	M	10th	White	Intellectual Disability	Other Health Impairment
Aaliyah	F	9th	Black	Intellectual Disability	
Jayla	F	10th	Black	Intellectual Disability	
Esteban	M	10th	Asian Pacific Islander	Intellectual Disability	Speech and Language Impairment

Table 1: Participant Demographic Information.

All participants followed the school's bell schedule to attend both core-content courses aligned to alternate achievement standards that were taught by special education teachers and elective courses with general education peers taught by general education teachers (e.g., Culinary). Mathematics instruction focused on computational fluency in a small group format but did not use a specific curriculum. The intervention occurred three to four times per week either in a teacher's lounge or conference room during each participant's math block for 15–35 min.

Interventionists

Two members of the research team rotated as the interventionist who had prior experience with MSBI and contributed to curriculum development (discussed further below). Before the study, interventionists reviewed procedural fidelity and role-played procedures for each condition. At the time, both interventionists were White females enrolled in a special education doctoral program, held current teaching certifications in special education, and had experience teaching students with ASD/ID as classroom teachers and researchers.

Experimental Design

A NCMB single-case experimental design (Harvey et al., 2004; Ledford & Gast, 2018; Slocum et al., 2022) measured the effect of MSBI on word problem solving behaviors. This variation of multiple-baseline design has garnered recent attention, with multiple author groups agreeing that it is a "practically useful and experimentally adequate design for answering some research questions" (Ledford, 2022, p. 662) as it still "provides verification with baseline logic, contrary to arguments made by prior researchers" (Smith et al., 2022, p. 667). Given our research questions, we used the logic of multiple baseline design and our own expertise in this content area to make design decisions that considered "the myriad of subtleties of research priorities and contextual factors" (Slocum et al., 2022, p. 681), rather than historical precedent for strict adherence to a specific set of rules (e.g., synchronized tiers; Ledford, 2022; Smith et al., 2022). As such, we aimed for a "balance between experimental rigor and innovativeness" (Harvey et al., 2004, p. 275).

Participants were assigned to tiers based on the sequential timing (i.e., first come first served) of when they turned in their consents and passed the prescreening tool (n = 4). Participants were randomly assigned to intervention start-point (i.e., baseline length) after screening using an online random generator. The online random generator randomized the number of baseline sessions between a range of possible baseline lengths: tier 1 between 3–4, tier 2 between 5–6, tier 3 between 7–8, and tier 4 between 9–10. The random generator selected 3, 6, 7, and 9 respectively.

There were three phases of intervention: ratio (unit 1), proportion (unit 2), and discrimination (unit 3). Participants completed one probe between each intervention phase wherein they solved one ratio and one proportion problem under

baseline conditions (see procedures for details). Intervention phase change decisions (e.g., Unit 1 to Unit 2) were based on participant data meeting the established mastery criteria for the primary dependent variable (described below). Based on Browder et al.'s (1986) guidelines, response-guided changes were considered when data had a stagnant or descending trend of three data points.

Measurement

The primary dependent variable was the number of critical independent problem-solving behaviors, measured by a researcher-created rubric that aligned with the problem-solving routine displayed on participant worksheets (Supplemental File 1). Although the problem-solving routine was shown as 8 steps on the student worksheets, Table 2 details how we measured 10 behaviors by splitting steps 5 and 7 into discrete behaviors for a total of 10 operationally defined and repeatedly measured problem-solving behaviors. Four behaviors were considered critical because they were required to arrive at the correct final answer (identified on Figure 1). The six non-critical behaviors were taught as a part of the problem-solving routine to support conceptual and procedural knowledge and measured to enable fine-grained tracking of participants success with MSBI as planned (see procedures for details). The mastery criteria for each unit were set a priori as three out of four critical problem-solving behaviors for two out of three sessions.

STEP	TASK ANALYSIS	OPERATIONAL DEFINITION	MEASUREMENT TYPE
1	Read to gather the facts	1. Read the word problem aloud.	NM
2	Circle the evidence to show what we know	2. Circle all quantities and corresponding nouns (for ratio, two numbers and at least two nouns; for proportion, three numbers and at least three nouns).	NC
3	Underline what we are investigating	3. Underline components of the question for what they are solving for: the label (i.e., noun).	NC
4	Discover the problem type	4a. Write the appropriate problem type on the line .	NC
		4b. Of the four choices, the correct problem type is indicated (e.g., draws circle, underlines, slash mark).	NC
5	Fill in the evidence on the correct schematic diagram	5a. Selects the correct schematic diagram.	NC
		5b. Fills out the correct, known quantities from the word problem on the schematic diagram.	C*
		5c. Fills out the correct, known labels from the word problem on the schematic diagram.	NC
6	Write the equation and solve problem	6. For ratio, answer is written three ways accurately with the unit quantities in the correct order; for proportion, equation is written with two equal sides with the variable.	C*
7	Report your findings	7a. Writes correct numerical response.	C*
		7b. Writes correct label.	C*
8	Check your sleuth skills	8. Checked their independent work alongside the interventionist.	NC

Table 2: Operational Definitions of Steps.

Note. C* = critical step; NC = non-critical step; NM = not measured.

Intervention: MSBI

This study was a feasibility test of three units of the Math Scene Investigators, a curriculum that uses MSBI to teach students to solve word problems. The curriculum was developed by the research team through an IEP-funded Early

Career Development and mentoring grant. All members of the research team contributed to the development of the materials, procedures, and protocols. Each unit had two introduction lessons, followed by daily lessons that consisted of one model problem, one guided practice problem, and one independent practice problem. Researchers delivered instruction using scripted lesson plans that incorporated the 16 elements of explicit instruction (Archer & Hughes, 2011). Lesson plan outlines can be found in Supplemental File 2.

Materials

Participant materials included worksheets, a choice board for schematic diagrams, a smartphone to access a calculator app, and self-management folders. Participants solved novel word problems each session that were written by researchers using Spooner et al.'s (2017) guideline and validated by a content area expert; no word problems were repeated throughout the study. Participants selected the schematic diagram from a laminated menu with the two options (ratio and proportion) and were given the corresponding schematic diagram as a 1" × 3" label sticker to place onto the worksheet. Vocabulary cards displayed a clear and concise definition and a visual representation with color coding for the following terms: ratio, proportion, equivalence, cross multiply, isolate the variable, and inverse operation (Supplemental File 3).

Baseline and Probe Sessions

Participants were given two worksheets (one for each problem type), a smartphone to access a calculator app, the schematic diagram laminated menu, and a pencil. The interventionist began by reading the word problem and provided the following statement: "Show me how to solve this problem. Remember, if you need anything read aloud again, you may ask me." As they independently completed both worksheets, the interventionist gave non-specific verbal praise (e.g., "You are working really hard!") but no corrective feedback.

Unit 1: Ratio

The first ratio lesson built conceptual understanding of ratios as a comparison of two quantities by explicitly teaching vocabulary (ratio and equivalence) and providing real world examples with concrete (two-colored counters), pictorial, and abstract representations of the quantities. Students were taught a "ratio rule" as a metacognitive strategy that paired precise language ("this is to that") with a hand gesture (Project Stair, 2019). A video model of the hand gesture for the ratio rule can be accessed at https://bit.ly/3SfVsGv. Interventionists modeled using the ratio rule to identify the relationship among quantities using planned examples and representing them in three ways (i.e., $3:1, 3 \text{ to } 1; \frac{3}{1}$). In the second ratio lesson, participants reviewed vocabulary and the ratio rule before the interventionist introduced the ratio schematic diagram and modeled how to use the problem-solving routine on the worksheet with a planned example.

After the two introduction lessons, interventionists followed scripted lessons with a model, guided practice, and independent practice format until the participant met mastery criteria. At the beginning of each session, participants reviewed their progress from the prior session and the goal they had set for this session (described in the self-monitoring routine below). Next, participants selected a location in their community from a picture menu (e.g., local pizza restaurant) and briefly discussed their experiences there. Two photographs from the selected location were used to stimulate the conversation as needed. Then the interventionist used a structured think aloud to model identifying the ratio relationship in the problem and writing the ratio three different ways by following the problem-solving routine on the worksheet.

Next, the interventionist and participant followed the problem-solving routine together with a second ratio word problem about the same community location. The interventionist gave either behavior specific praise or corrective feedback (i.e., error correct with model-retest) for each problem-solving behavior using clear, concise, and consistent language (Powell & Fuchs, 2018). Finally, participants independently practiced the problem-solving routine with a third

ratio word problem about their selected community location. The interventionist provided no behavior-specific praise or corrective feedback while participants solved the third problem independently. After finishing the independent practice problem, participants checked their work and engaged in the self-monitoring routine (described below).

Unit 2: Proportions

The first proportion lesson built conceptual understanding of proportions as two equivalent ratios. Participants were explicitly taught vocabulary (ratio, proportion, and equivalence/equivalent) and the new vocabulary cards were displayed for participants. The metacognitive strategy was called the "proportion rule," which expanded on the "ratio rule" to reinforce their relationship: "this is to that is the same as this is to that." In the first lesson, participants identified equivalent relationships in real-world examples of proportions where all quantities are known (e.g., "This is to that is the same as this is to that; 1 avocado to 3 people is the same as 3 avocados to 9 people."). Two-sided counters were used to reinforce conceptual understanding. A video of the proportion rule can be accessed at bit.ly/4kk2pSY.

In the second proportion lesson, participants began by reviewing concepts from the first lesson, learned new vocabulary (cross multiply, isolate the variable, and inverse operation), and practiced using the proportion rule to model scenarios with equivalent ratios. Next, the interventionist introduced the proportion schematic diagram and modeled how to solve a proportion problem with an unknown quantity (e.g., "This is to that is the same as this is to that; 3 peppers to 1 tortilla is the same as how many peppers to 2 tortillas?"). The interventionist modeled how to apply the inverse operation to isolate the variable to solve for the unknown using cross multiplication. After these two introductory lessons, they followed the same model, guided practice, and independent practice format using novel word problems until participants met mastery criteria.

Unit 3: Discrimination

The first two discrimination lessons focused on applying the metacognitive strategy to identify the key features of each schema within word problems using multiple exemplar training. The interventionist began by using a T-chart sorting activity with a think aloud of the key features of each problem type (e.g., rule, schematic diagram). Next, participants sorted word problems independently and added information from the problem into the rule to demonstrate conceptual understanding (e.g., "comparing this to that, 4 tacos to 1 person"). Finally, participants solved one ratio and one proportion word problem with the interventionist (i.e., guided practice) and two novel problems independently. In subsequent discrimination lessons, participants solved two problems independently (one of each type), checked their work, and engaged in the self-monitoring routine until they reached mastery criteria for both problem types.

Self-Monitoring Routine

After solving the independent practice problem(s), participants checked their work with assistance from the interventionist. They self-monitored accuracy of each problem solving behavior directly on the worksheet by giving themselves a check mark for those completed independently correct. Any errors were discussed and revised using a green pen. Participants recorded the total number of check marks (i.e., independent correct behaviors) and used their self-reflection folder to: (a) self-record and self-graph their score from the independent practice problem(s), (b) self-evaluate whether or not they met the goal they had set, (c) set a goal for the next session, and (d) self-reflect on their problem solving experience using two Likert scales represented with emojis and by circling icons representing the components of MSBI they enjoyed (e.g., goal setting, checking off their work on the task analysis)

Interobserver Agreement (IOA) and Procedural Fidelity (PF)

To reduce bias, we used an online random number generator to identify a minimum of 30% of sessions for each participant for IOA and PF before the study began (Ledford et al., 2023). A graduate research assistant not involved in

data collection was trained to use the coding manual to independently score worksheets. Disagreements were discussed in weekly meetings. IOA was calculated using point-by-point agreement (Ledford & Gast, 2018), with an overall average of 96% agreement in baseline and probes (range of 90–100%), 90% for intervention sessions (range of 80–100%), and 90% for discrimination (range of 60–100%) across all scored sessions and interventionist. A PF checklist was used to measure critical elements of the intervention (Supplemental File 4). A total of 41% of all sessions were scored for PF for both interventionist across all sessions; PF was overall average of 98% for baseline and probes (range of 80–100%), 96% for intervention sessions (range of 83–100%), and 99% for discrimination (range of 94–100%). Supplemental File 5 details IOA and PF broken down by Unit across each participant.

Data Analysis

Following Kratochwill et al. (2022) recommendations for NCMB data presentation, we chose to display the data as if it were collected in real-time by aligning sessions across each tier (participants) to enable vertical analysis (Figure 1). Dates on the x-axis indicate when each participant's data collection began. To address long-standing concerns in SCED regarding interrater agreement in visual analysis, we followed Wolfe et al.'s (2019) protocol for systematic visual analysis
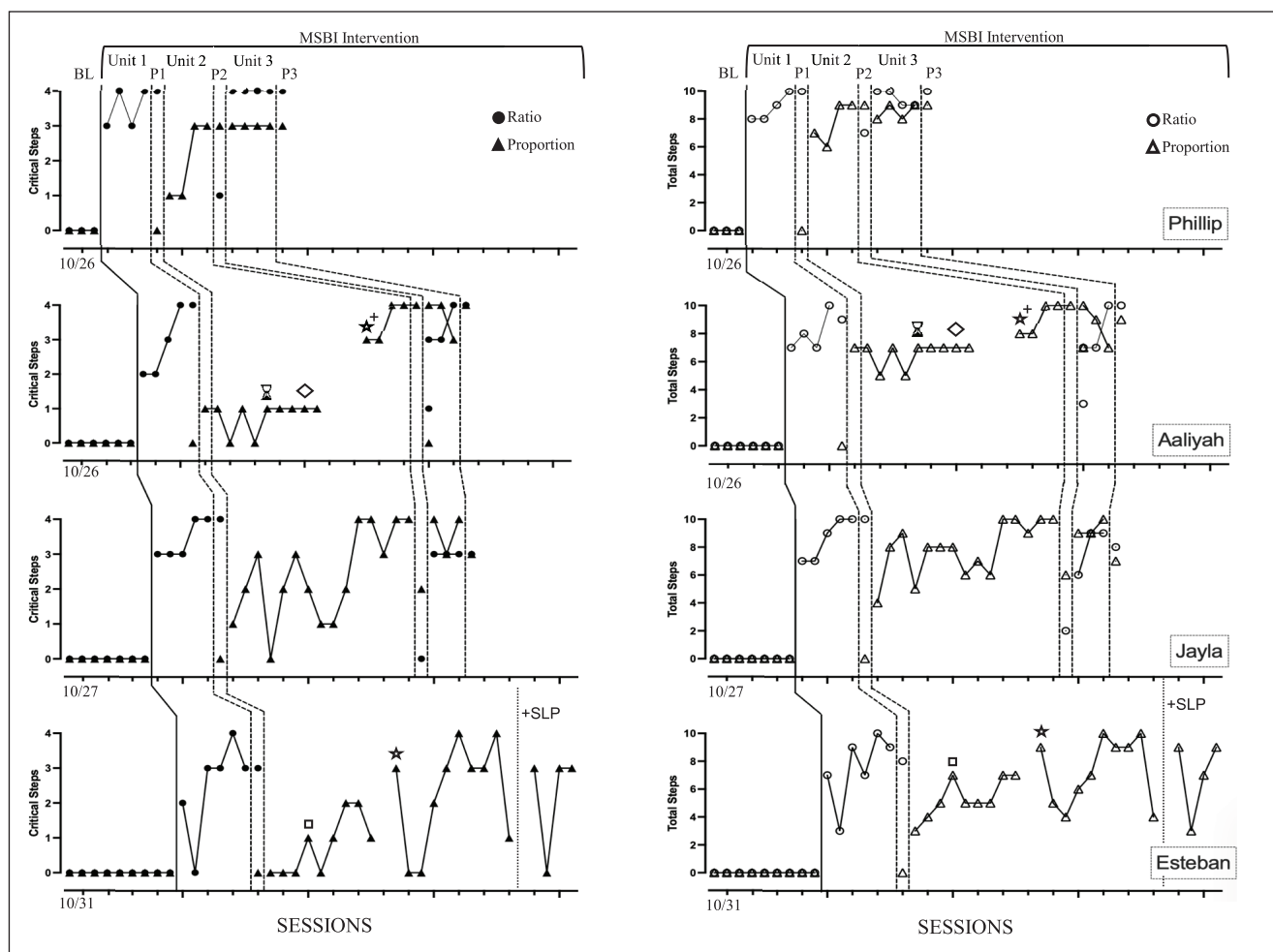

Figure 1: Graph of problem solving behaviors.

Note. response-guided changes: hourglass = self-monitoring; rhombus = remove model problem; star+ = massed trial with extended task analysis; rectangle = read aloud, dictation, and highlighters; star = massed trial.

of level, trend, variability, immediacy, overlap, and consistency. Separate visual analysis for ratio and proportion enabled interpretation of effects within and between phases.

Social Validity

After the intervention concluded, interventionists engaged participants in social validity interviews using a semi-structured protocol of six open-ended questions and five Likert rating scale questions with emojis and text to represent response options (Supplemental File 6). All questions and response options were read aloud and displayed to participants one at a time. Pictures of study materials (e.g., calculator, hand motion rules, schematic diagrams, task analysis) were displayed on a printed matrix as a visual reference. Participants could request for a break between open ended and Likert scale questions. Social validity data was analyzed using content-analysis.

RESULTS

Figure 1 displays a graph of correct problem-solving behaviors during independent practice problems. The left graph contains student performance on the critical steps for both ratios (closed circlers) and proportions (closed triangles) while the graph on the right shows student performance on total steps with open circles (ratios) and open triangles (proportions).

Research Question 1: Effect of MSBI on Problem-Solving Behaviors

Unit 1: Ratios

Baseline data for ratio problems were stable with zero trend across all four participants. Each participant demonstrated an immediacy of effect for ratio problems. The data paths display a consistent level change and an accelerating trend replicated across participants. Esteban's data were more variable when compared to the other participants, with one overlapping data point with baseline. All participants reached mastery criteria after six sessions and maintained performance during the first probe. Visual analysis of Figure 1 identified a functional relation with four demonstrations of effect for ratio problems.

Unit 2: Proportions

Baseline data for proportion problems were stable with zero trend across all four participants. Visual analysis indicated an immediacy of effect for three of four participants. As anticipated, the initial level change for proportion problems was smaller and data were more variable compared to ratio problems. Phillip's scores immediately increased from 0 to 1 and continued upward until he met mastery in the fourth session with no overlap between baseline, probe 1, and Unit 2. Although Aaliyah's performance jumped after beginning Unit 2, the first five sessions were variable, and two data points overlapped with baseline. Following response-guided changes (described below) she met mastery criteria after 15 sessions. Jayla's data immediately increased in level after beginning Unit 2, with only one overlapping data point. She also met mastery criteria after 15 sessions. Esteban's data did not demonstrate immediacy of effect, with 29% overlap between baseline/probe data points and proportion data points. Esteban did meet mastery criteria after a series of response-guided changes (described below). Visual analysis of Figure 1 identified a functional relation with three demonstrations of effect for proportion problems.

Unit 3: Discrimination Training

Three participants (Phillip, Aaliyah, and Jayla) completed probe 2 before moving into the discrimination unit. No participants maintained skills for ratio problems and only Phillip maintained skills for proportion problems. This data met our expectation that participants would overgeneralize, with all participants demonstrating higher performance in the most recently taught problem type (proportion) despite having previously demonstrated mastery of ratio problems.

All three participants maintained mastery in ratio and proportions after three (Jayla, Aaliyah) and four (Phillip) discrimination sessions and maintenance in probe 3.

Research Question 2: Impact of Response-Guided Changes

As anticipated, response-guided changes to MSBI were necessary for two participants. Symbols in Figure 1 identify the onset of each change, which are described further below.

Aaliyah

Both interventionists observed a change in Aaliyah's affect and stamina (e.g., slow and delayed engagement with materials, putting her head down) once they began Unit 2. The stagnant data pattern was attributed to fatigue given the increased duration of proportion intervention sessions compared to ratio intervention sessions. The team hypothesized a motivation-related performance deficit (i.e., MSBI was not adequately supporting engagement). To increase engagement, we adjusted the self-monitoring system to focus on fluency instead of just accuracy (hourglass in Figure 1).

While this change appeared to increase Aaliyah's motivation and desire to engage in the lessons (e.g., actively engaging with study materials, verbally expressing her excitement to work with interventionists), it did not impact the measured problem-solving behaviors or decrease the duration of lessons. Still working from the hypothesis that the duration of the session was the barrier to progress, the research team reduced the task demand and duration of the session by only presenting guided and independent practice problems (rhombus in Figure 1). After two sessions with no measurable improvement, we concluded there must also be a skill deficit.

Step-level analysis of her data across sessions revealed consistent procedural errors in step 6 ("write and solve the problem"), which prevented accuracy in step 7 ("report your findings"). Before interventionists could address this perceived performance deficit, Aaliyah was absent for an extended time due to an illness, as indicated on the graph with hash marks starting after session 10. Once Aaliyah returned to school, the interventionists gave her isolated opportunities to practice using an expanded task analysis for step 6: "Write an equation and solve the problem," breaking this into four discrete behaviors paired with visuals: (1) set up your equation, (2) simplify your equation, (3) isolate the variable using the inverse operation, and (4) write your answer with the label (marked with a star and superscript plus sign in Figure 1). She chose to continue using the extended task analysis and met mastery criteria after five sessions.

Esteban

Esteban was not able to read the problem independently and would not ask interventionists to re-read it. Therefore, three changes were made beginning in the fourth session (marked by a rectangle in Figure 1). First, he was taught to use a text-to-speech app on his phone as an alternate means of representation and was told he could dictate to the interventionist for the "report your findings" segment of the worksheet. He was also taught to use a highlighter to make the important information in the word problem more salient (e.g., 80 m; 10 s). He began accurately transferring information from the problem to the schematic diagram, leading to a slight upward trend over the next six intervention sessions.

Similarly to Aaliyah, Esteban's data indicated difficulty with using cross multiplication to solve for the missing variable. After massed trial practice using the extended task analysis, he was able to correctly write an equation representing the proportional relationship. We hypothesized variability suggested a lack of conceptual understanding. We increased the efficiency and intensity of instruction by using a system of least prompts (verbal, specific-verbal, model-retest) during independent practice (marked by phase line in Figure 1). Interventionists anecdotally observed an improvement in his performance, but his data remained variable. Esteban requested to be done with the research and stay within his regularly schedule mathematics class rather than continue into Unit 3. He disclosed various changes occurring within his regularly scheduled mathematics class (e.g., peer mentors) and that he did not want to miss those activities.

Social Validity

Social validity interviews ranged from 15–21 minutes (average of 18 minutes). All participants liked setting daily goals, choosing word problem locations, checking their work with the interventionists, and graphing their progress. Phillip shared that he liked checking his work because, "I like seeing the results" and Jayla recalled that she enjoyed graphing her scores because "I like seeing how I'm doing with my math. I like seeing my progress." Participants identified which mathematics tools were the most helpful during their learning: schematic diagrams (all), the extended checklist (Aaliyah), and the text-to-speech app (Esteban). Esteban shared that he was using the text-to-speech app in other settings outside of the study, and that his caregivers planned to purchase the paid version of the app. Three of the four participants expressed a desire for their mathematics teacher to teach them the way they learned it with us in the research study.

DISCUSSION

This study reports findings from a feasibility test of three units from Math Scene Investigators, a curriculum that uses MSBI to teach multiplicative word problem solving. Using Wolfe et al.'s (2019) systematic protocol for visual analysis, we concluded there was a functional relation between MSBI and problem-solving behaviors for both ratio and proportion problems.

Alignment with Anticipated Data Patterns

We applied this logic of systematic replication in SCED (Travers et al., 2016) by assuming that if our participant characteristics, task demands, and instructional supports were similar to those in prior published studies then MSBI *as planned* would at minimum result in an increase in measured behaviors for the first problem type (ratio) over baseline levels for all participants within the first few sessions. Each of our participants met this expectation in the first unit on ratios as demonstrated by an immediate change in level and trend replicated across all four participants. Regarding the proportions, we anticipated all participants would need an increase in dosage of the intervention to reach criteria due to the more complex syntax and procedures (i.e., cross multiplication). This expectation was also met, as participants required more instructional sessions to meet criteria for proportions (range 15–25) over ratios (range 4–6).

Finally, our expectations were met in terms of participants requiring explicit discrimination training (Unit 3), given that prior MSBI studies that targeted more than one problem type documented this need (e.g., Browder et al., 2018; Cox et al., 2021; Root et al., 2020; Saunders, 2014). While schema instruction is intentionally designed to support generalization and transfer (Fuchs et al., 2004; Spooner et al., 2017), probes showed our participants *overgeneralized*, meaning they applied the most recently taught strategy to all problems until given explicit instruction and deliberate practice with planned examples.

Post-Hoc Exploratory Analyses of Response-Guided Changes

Whereas hypothesizing after results are known (i.e., Harking) is a questionable research practice, Hollenbeck et al. (2016) argue there is value in *transparently* reporting new understandings or hypotheses derived from post-hoc analyses when they are "driven by informed reconsiderations that prompt interesting questions…particularly in high-stake contexts where data are resource-intensive to obtain" (p. 11). They refer to this as "Tharking," defined as "clearly and transparently reporting new hypotheses that were derived from post-hoc results in the Discussion section of an article", with emphasis on how (transparently) and where (in the Discussion section) this takes place (p. 11). In this study, post-hoc exploratory analyses informed our understanding of the effectiveness of MSBI *as planned* as well as the need for and outcome of response-guided changes. To guard against Harking, the content of the results section reflects our rationale for and

interpretation of response-guided changes as they were made during the study. Here we aim to transparently discuss how our post-hoc analyses impacted our interpretation of participants' need for and response to response-guided changes.

First, it is important to reiterate that our a priori rules for making response-guided changes were based on long-established recommendations for making data-based decisions (Browder et al., 1986) and those reported in prior published MSBI studies (see Cox et al., 2024). We attributed variability to a skill deficit and zero to minimal trends to a performance deficit (e.g., can-do vs. won't do; Codding et al., 2017). The UDL framework was then used to consider the barriers in the task and corresponding solutions based on that categorization. For example, Esteban's stagnant data pattern in the beginning of Unit 2 was considered a skill deficit; he was unable to read the problem and therefore made errors in filling out the schematic diagram. MSBI is designed to address this barrier by having the student ask for things to be read aloud by an adult or peer (Spooner et al., 2017), but this alternate mode of representation presented a barrier to engagement for Esteban because he was self-conscious about asking for this support. Therefore, the first response-guided change made for Esteban was a self-directed option for accessing the problem (i.e., text-to-speech app) that minimized threats and distractions. As reflected in the results above, access to the problem was not considered as a barrier after he began using the app.

At the conclusion of data collection, a member of the research team conducted an error analysis of all worksheets for Aaliyah, Jayla, and Esteban. They discovered direct alignment between dips in the dependent variable and independent practice problems that had multi-word labels (e.g., ice cream cone and bowl of ice cream), units (e.g., cups of rice, ounces of water), or had labels beginning with the same letter (e.g., apple and avocado). We did not anticipate language to be a barrier within the task because we followed the same guidelines as prior MSBI studies and had our problems validated by an expert in mathematics education. We concluded Spooner et al.'s (2017) guidelines for writing word problems may not sufficiently encompass the universe of variables that influence cognitive accessibility for students with ASD/ID, particularly those with limited literacy skills who must rely on listening comprehension.

This discovery prompted post-hoc analysis of both the need for and outcomes of response-guided changes as we generated new exploratory hypotheses regarding *for whom* and *under what conditions* MSBI "works." Executive functioning emerged as a potential barrier, specifically demands on working memory, impulse control, and cognitive flexibility as these are three components of executive functioning that Purpura et al. (2017) identified as being distinctly related to development of mathematics skills. Our exploratory post-hoc analyses found evidence of executive dysfunction during Unit 2 for Aaliyah and Esteban. For example, they had more difficulty staying on task (inhibition control), transferring quantities and labels from the problem into the schematic diagram (cognitive flexibility, working memory), and executing the multi-step cross multiplication procedure (working memory) than in Unit 1.

Our inductive approach led us to consider whether cognitive load theory may give a helpful theoretical and empirical basis for proactively identifying barriers in the learning environment and making response-guided changes to MSBI. Cognitive load theory is conceptually aligned with schema-induction theory, which focuses on the development and activation of mental models and schemas. Cognitive load theory assumes the finite capacity of an individual's working memory is a barrier to storage of new information in long-term memory due to an intricate relationship between working memory and three types of cognitive load (Sweller, 2010). Intrinsic cognitive load represents the inherent mental effort required by the task and extraneous cognitive load refers to the unnecessary mental effort caused by factors unrelated to the learning task. For example, intrinsic cognitive load of a word problem solving task refers to mental effort required to process and understand the information presented in the problem itself. Extraneous cognitive load results from information that must be processed but is not directly relevant, such as extraneous information in a word problem. High intrinsic and extrinsic cognitive loads reduce germane cognitive load, which is the mental effort dedicated to constructing meaningful relationships between new information and existing knowledge. Both

cognitive load theory and schema induction theory aim to optimize learning by activating relevant schemas and provide complementary frameworks for reducing reliance on working memory by retrieving and activating existing schemas from long-term memory.

We hypothesize the variability in language demands and lexical ambiguity of the proportion word problems likely increased their intrinsic cognitive load (Gupta & Zheng, 2020), particularly for Esteban, the only participant with a secondary language impairment. Relatedly, the expanded task analysis and massed trial practice of the cross-multiplication procedure may have decreased extraneous load by reducing demands on working memory. Given that the contribution of working memory to individual differences in word problem solving for students with learning disabilities is well documented (Fuchs et al., 2020; Lee et al., 2009; Swanson, 2011) and individuals with ID perform significantly lower on executive functioning tasks compared to typically developing peers (Spaniol & Danielsson, 2022), this is an area that warrants further inquiry.

LIMITATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

Establishing guidelines for making response-guided changes to MSBI (i.e., data-based decision making) that are specific to the learning profile of students with ASD/ID would be useful to both researchers and practitioners. Although authors of extant MSBI studies have consistently reported the response-guided changes and marked their onset using symbols on their graphs, they have not reported the decision-rule system used to make such changes and our exploratory post-hoc analysis pointed out limitations of our approach. While our findings may provide preliminary insight into cognitive skills that may influence success in word problem solving (i.e., executive functioning), our response-guided decisions were made using a combination of prior research and professional expertise. Furthermore, the one-on-one instructional pull-out format with researchers as interventionists does not reflect natural teaching conditions. This impacts external validity and generalizability of findings, as practitioners generally would not have the level of expertise, control, and resources when implementing Math Scene Investigators in their own classrooms. Additionally, this instructional format required students to be pulled from their regularly schedule mathematics class, reducing the amount of time they were engaging in mathematics in the naturalistic classroom based setting with their peer mentors. As Esteban disclosed, this had a direct influence on his participation and removal from the study (i.e., he wanted to be with his peer mentor).

Also, we want to acknowledge the limitations of our social validity data, as it was only collected post-intervention (Schwartz & Baer, 1991). Future research in this area should explore ways to embed more opportunities to understand participant perspectives on the goals, procedures, and outcomes of MSBI. Finally, some participants had prior experience with MSBI, which may have influenced their response to the intervention despite not showing generalization on the screening. Studies with natural interventionists (e.g., teachers, paraprofessionals, peers) under realistic conditions (e.g., small groups within a classroom environment) are needed to fully understand the conditions under which MSBI "works," both *as planned* and following response-guided changes. Our new hypotheses regarding the utility of cognitive load theory to ameliorate barriers in the learning task resulting from variability in executive functioning warrants inquiry.

Implications for Practice

Word problem solving is one way for students ASD/ID to apply mathematical knowledge and skills aligned with situations they could encounter in everyday life. Additionally, while the alternate achievement standards are reduced in depth, breadth, and complexity when compared to the grade-level standards, they still emphasis mathematical word problem solving. Our study highlights the effectiveness of MSBI for teaching word problem-solving to students with ID while emphasizing the need for proactive and reactive interventions tailored to individual needs. Practitioners can adapt mathematical instruction to support the cognitive load of their learners by strategically minimizing intrinsic cognitive

load (i.e., aligning instruction to students learning level, iteratively introducing new concepts, and explicitly teaching important vocabulary) and extraneous cognitive load (i.e., limiting extraneous information, chunking instruction, limiting distractions) to support student learning.

Practitioners should proactively follow a universal design approach using Spooner et al.'s (2017) model for designing the instruction and materials, and reactively use student data to address specific difficulties encountered during instruction. To do so, discrete problem-solving behaviors must be operationally defined and repeatedly measured to identify specific difficulties (e.g., cross multiplication, information transfer). Observed difficulties should be practiced in isolation using massed trials with frequent opportunities to respond.

SUPPLEMENTARY FILES

- **Supplemental File 1:** Worksheet & Operational Definitions. https://doi.org/10.25894/rise.2778.s203
- **Supplemental File 2:** Lesson Outlines. 10.25894/rise.2778.s204
- **Supplemental File 3:** Vocabulary Cards. 10.25894/rise.2778.s205
- **Supplemental File 4:** PF Checklists. 10.25894/rise.2778.s206
- **Supplemental File 5:** IOA & PF. 10.25894/rise.2778.s207
- **Supplemental File 6:** Social Validity. 10.25894/rise.2778.s208

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Deidre Gilley orcid.org/0000-0002-1216-5967
Educational Psychology, Baylor, US

Jenny Root orcid.org/0000-0003-4682-0832
Florida State University, US

Alicia Saunders orcid.org/0000-0002-1313-5719
School of Teacher Education, Florida State University, US

Sarah Cox orcid.org/0000-0002-2074-8839
Michigan State University, US

Danielle Morsching orcid.org/0009-0004-7636-051X
Georgia Southern University, US

AUTHOR CONTRIBUTIONS

Deidre Gilley: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Resources, Supervision, Validation, Visualization, Writing – Original Draft, Writing – Reviewing & Editing. **Jenny Root**: Conceptualization, Formal Analysis, Funding Acquisition, Methodology, Project Administration, Resources, Supervision, Visualization, Writing – Original Draft, Writing – Reviewing & Editing. **Alicia Saunders**: Conceptualization, Resources, Visualization, Writing – Original Draft, Writing – Reviewing & Editing. **Sarah Cox**: Conceptualization, Formal Analysis, Resources, Visualization, Writing – Original Draft, Writing – Reviewing & Editing. **Danielle Morsching**: Investigation, Resources, Validation, Visualization, Writing – Original Draft, Writing – Reviewing & Editing.

REFERENCES

Archer, A. L., & Hughes, C. A. (2011). *Explicit instruction: Effective and efficient.* Guilford Press.

Browder, D. M., Liberty, K., Heller, M., & D'Huyvetters, K. K. (1986). Self-management by teachers: Improving instructional decision making. *Professional School Psychology*, 1(3), 165.

Browder, D. M., Spooner, F., Lo, Y., Saunders, A. F., Root, J. R., Ley Davis, L., & Brosh, C. R. (2018). Teaching students with moderate intellectual disability to solve word problems. *The Journal of Special Education*, 51(4), 222–235. https://doi.org/gcvm9k

Center for Applied Special Technology (CAST; 2018). *Universal Design for Learning guidelines version 2.2.* http://udlguidelines.cast.org

Clausen, A. M., Tapp, M. C., Pennington, R. C., Spooner, F., & Teasdell, A. (2021). A systematic review of modified schema-based instruction for teaching students with moderate and severe disabilities to solve mathematical word problems. *Research and Practice for Persons with Severe Disabilities*, 46(2), 94–107. https://doi.org/h2xd

Codding, R. S., Volpe, R. J., & Poncy, B. C. (2017). *Effective math interventions: A guide to improving whole-number knowledge.* Guilford.

Cox, S. K., Root, J. R., Goetz, K., & Taylor, K. (2021). Modified schema-based instruction to encourage mathematical practice use for a student with autism spectrum disorder. *Education and Training in Autism and Developmental Disabilities, 56*(2), 190–204. http://www.daddcec.com/uploads/2/5/2/0/2520220/etadd_56_2_june_ii.pdf

Cox, S. K., Root, J. R., McConomy, A., & Davis, K. (2024). "For whom" and "under what conditions" is MSBI effective? A conceptual replication with high school students with autism. *Exceptional Children*, 90(4), 361–381.

Fuchs, L. S., Fuchs, D., Prentice, K., Hamlett, C. L., Finelli, R., & Courey, S. J. (2004). Enhancing mathematical problem solving among third-grade students with schema-based instruction. *Journal of Educational Psychology*, 96(4), 635. https://doi.org/dpzbr5

Fuchs, L., Fuchs, D., Seethaler, P. M., & Barnes, M. A. (2020). Addressing the role of working memory in mathematical word-problem solving when designing intervention for struggling learners. *ZDM*, 52(1), 87–96. https://doi.org/gk355g

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15(1), 1–38.

Gupta, U., & Zheng, R. (2020). Cognitive load in solving mathematics problems: Validating the role of motivation and the interaction among prior knowledge, worked examples, and task difficulty. *European Journal of STEM Education*, 5(1), 05. https://doi.org/mstq

Harvey, M. T., May, M. E., & Kennedy, C. H. (2004). Nonconcurrent multiple baseline designs and the evaluation of educational systems. *Journal of Behavioral Education*, 13(4), 267–276. https://doi.org/dwbk6g

Jitendra, A. K., & Star, J. R. (2012). An exploratory study contrasting high- and low-achieving students' percent word problem solving. *Learning and Individual Differences, 22*(1), 151–158. https://doi.org/frpwcs

Kratochwill, T. R., Levin, J. R., Morin, K. L., & Lindström, E. R. (2022). Examining and enhancing the methodological quality of nonconcurrent multiple-baseline designs. *Perspectives on Behavior Science*, 45(3), 651–660. https://doi.org/hxpv

Ledford, J. R. (2022). Concurrence on nonconcurrence in multiple-baseline designs: A commentary on Slocum et al. (2022). *Perspectives on Behavior Science*, 45(3), 661–666. https://doi.org/mstt

Ledford, J. R., & Gast, D. L. (Eds.). (2018). *Single case research methodology: Applications in special education and behavioral sciences* (3rd ed.). Routledge.

Ledford, J. R., Lambert, J. M., Pustejovsky, J. E., Zimmerman, K. N., Hollins, N., & Barton, E. E. (2023). Single-case-design research in special education: Next-generation guidelines and considerations. *Exceptional Children*, 89(4), 379–396. https://doi.org/grsphf

Lee, K., Ng, E. L., & Ng, S. F. (2009). The contributions of working memory and executive functioning to problem representation and solution generation in algebraic word problems. *Journal of Educational Psychology*, 101(2), 373–387. https://doi.org/cw3swv

Lein, A. E., Jitendra, A. K., & Harwell, M. R. (2020). Effectiveness of mathematical word problem solving interventions for students with learning disabilities and/or mathematics difficulties: A meta-analysis. *Journal of Educational Psychology*, *112*(7), 1388–1408. https://doi.org/gkqg

Powell, S. R., & Fuchs, L. S. (2018). Effective word-problem instruction: Using schemas to facilitate mathematical reasoning. *TEACHING Exceptional Children*, *51*(1), 31–42. https://doi.org/gf86qb

Project Stair. (2019). *Schema gestures || word problems* [Video]. YouTube. https://www.youtube.com/watch?v=siPbBJ2NyZk

Purpura, D. J., Schmitt, S. A., & Ganley, C. M. (2017). Foundations of mathematics and literacy: The role of executive functioning components. *Journal of Experimental Child Psychology*, *153*, 15–34.

Root, J. R., Cox, S. K., Saunders, A., & Gilley, D. (2020). Applying the universal design for learning framework to mathematics instruction for learners with extensive support needs. *Remedial and Special Education, 41*(4), 194–206. https://doi.org/ggqmxq

Root, J. R., Ingelin, B., & Cox, S. K. (2021). Teaching mathematical word problem solving to students with autism spectrum disorder: A best-evidence synthesis. *Education and Training in Autism and Developmental Disabilities*, *56*(4), 420–436.

Root, J. R., Saunders, A., Cox, S. K., Gilley, D., & Clausen, A. (2022). Teaching word problem solving to students with autism and intellectual disability. *TEACHING Exceptional Children*, 4005992211168.

Saunders, A. F. (2014). *Effects of schema-based instruction delivered through computer-based video instruction on mathematical word problem solving of students with autism spectrum disorder and moderate intellectual disability*. (Order No. 3636163). [Doctoral dissertation, University of North Carolina at Charlotte]. ProQuest Dissertations Publishing.

Slocum, T. A., Pinkelman, S. E., Joslyn, P. R., & Nichols, B. (2022). Threats to internal validity in multiple-baseline design variations. *Perspectives on Behavior Science*, *45*(3), 619–638. https://doi.org/mfb3

Smith, S. W., Kronfli, F. R., & Vollmer, T. R. (2022). Commentary on Slocum et al. (2022): Additional considerations for evaluating experimental control. *Perspectives on Behavior Science*, *45*(3), 667–679. https://doi.org/mstx

Spaniol, M., & Danielsson, H. (2022). A meta-analysis of the executive function components inhibition, shifting, and attention in intellectual disabilities. *Journal of Intellectual Disability Research*, *66*(1–2), 9–31.

Spooner, F., Saunders, A., Root, J., & Brosh, C. (2017). Promoting access to common core mathematics for students with severe disabilities through mathematical problem solving. *Research and Practice for Persons with Severe Disabilities*, *42*(3), 171–186. https://doi.org/gbtn2w

Swanson, H. (2011). Working memory, attention, and mathematical problem solving: A longitudinal study of elementary school children. *Journal of Educational Psychology*, *103*(4), 821–837. https://doi.org/dz22wf

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, *22*(2), 123–138. https://doi.org/dmvdmq

Tincani, M., & Travers, J. (2022). Questionable research practices in single-case experimental designs: Examples and possible solutions. In *Avoiding questionable research practices in applied psychology* (pp. 269–285). Cham: Springer International Publishing.

Travers, J. C., Cook, B. G., Therrien, W. J., & Coyne, M. D. (2016). Replication research and special education. *Remedial and Special Education*, *37*(4), 195–204.

Wolfe, K., Barton, E. E., & Meadan, H. (2019). Systematic protocols for the visual analysis of single-case research data. *Behavior Analysis in Practice*, *12*(2), 491–502. https://doi.org/gnm4x2

aperio