Examination of a Modified Incremental Rehearsal Approach to Explore Causal Mechanisms

Kathrin E. Maki · Mary Elizabeth Moody · Siera L. Cullins · Taylor L. Griffin

Accepted: 14 June 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
Incremental rehearsal (IR) has consistently been shown to improve students’ math fact retention and fluency (Maki et al., Journal of Behavioral Education 30:534–558, 2021). However, less is known about how intervention modifications may support longer-term skill maintenance. The purpose of this study was to compare traditional IR with a modified IR (shuffle IR; ShIR) in which known multiplication facts were shuffled between sequences using a cumulative acquisition design with six fourth- and fifth-grade students. All participants retained and maintained more facts in IR and ShIR compared to a control condition. However, IR or ShIR did not consistently result in greater retention than the other, with three students demonstrating greater retention in the IR condition and three students demonstrating greater retention in the ShIR condition. Most participants demonstrated greater fact maintenance in the ShIR condition than in the IR condition. All participants made fewer intervention session errors in the condition in which they retained more multiplication facts.

Keywords Math intervention · Incremental rehearsal · Single-case design

Introduction
Proficiency in early math skills is strongly related to later math development and proficiency in other academic areas (e.g., reading; Bailey et al., 2014; Duncan et al., 2007; Purpura et al., 2017). Knowledge of foundational math concepts (e.g., math facts) has been shown to predict acquisition of more complex skills throughout development across academic domains (Bailey et al., 2014; Cameron et al., 2019; Watts et al., 2014, 2015). However, in 2019, only 41% of fourth grade students demonstrated proficiency in math on the National Assessment of Educational Progress (NAEP; Nation’s Report

* Kathrin E. Maki
kathrin.maki@coe.ufl.edu

1 Department of Special Education, School Psychology, and Early Childhood Studies, University of Florida, Gainesville, FL, USA

Published online: 28 July 2022
Card, 2019). Students experiencing academic difficulty in schools were more likely to engage in problem behaviors (Wang & Fredricks, 2014) and at greater risk for dropping out of school (Fan & Wolters, 2014). Conversely, students demonstrating early math proficiency, including with math facts, have been shown to complete higher level math courses in high school, resulting in increased likelihood of college attendance and graduation and higher lifetime earnings (Lee, 2012). These significant outcomes associated with math proficiency underscore the need for effective math intervention supports in schools to ensure that students demonstrate necessary foundational skills such as math fact knowledge to support successful engagement in complex academic tasks as well as later independent functioning. Targeting early math skills such as math facts is critical to support equitable opportunities for all students in school and in future careers (Van-DerHeyden & Codding, 2020). The purpose of this study was to compare the effectiveness of two drill interventions on multiplication fact retention (i.e., next session correct responding) and maintenance (correct responding at one-, two-, and three weeks post-intervention) and to compare response error rate within the interventions.

**Math Fact Intervention**

Many students with math learning disabilities (MLD) and math difficulties (MD) struggle with math fact knowledge (e.g., 5 + 8, 4 × 3; Calhoon et al., 2007; Nelson et al., 2022), which can interfere with development of higher order math skills (e.g., algebraic skills) and complex problem solving because cognitive resources must be devoted to the lower order math skills (i.e., math fact; McVancel et al., 2018). Mastery and automatic retrieval of math facts reduces cognitive load and allows for devotion of cognitive resources to more complex tasks (e.g., algebraic operations), which, in turn, supports overall math proficiency (Woodward, 2006). Therefore, math facts represent an important intervention target for students with MLD and MD to support overall math proficiency.

Math interventions should be targeted to students’ needs and should incorporate instructional strategies that support automaticity, or quick, effortless, and low error responding (Burns et al., 2015; Delazer et al., 2005). Interventions should incorporate modeling, explicit instruction, corrective feedback, novel practice opportunities, independent practice, and timing in response to acquisition and/or fluency needs (Haring & Eaton, 1978; Ardoin & Daley, 2007; Maki et al., 2021; McKevett & Codding, 2021). Incorporating these instructional strategies has been shown to improve students’ basic academic skills (Ardoin et al., 2018), including math fact acquisition and fluency (Burns et al., 2016). Moreover, matching math fact interventions to students’ specific skill difficulties and level of skill development has been shown to result in greater short-term math fact retention and fluency than when interventions were not correctly matched (Maki et al., 2021).
Incremental Rehearsal

Drill interventions are one approach commonly used to target foundational skills (e.g., letter name/sound knowledge, vocabulary), including math facts, and have been shown to improve higher order skills when they incorporate effective instructional components such as repeated skill practice, modeling, and corrective feedback (Burns, 2005). Incremental rehearsal (IR; Tucker, 1989) is one drill intervention that can be used to target math fact acquisition and fluency. IR intersperses known and unknown stimuli, with the number of known stimuli presented increasing throughout the fact presentation sequence (IR sequence presented in Fig. 1). IR has demonstrated effectiveness for basic skill acquisition generally and for math facts specifically (Burns et al., 2010; Codding et al., 2009, 2011; Maki et al., 2021). However, most studies examining IR’s effectiveness have focused on shorter-term skill retention and fluency rather than longer-term skill maintenance (e.g., Burns, 2005; McVancel et al., 2018).

IR incorporates several instructional features supporting its effectiveness. First, IR provides students with a high number of opportunities to respond (OTR; Sadokierski & Burns, 2008) to the unknown stimulus by repeatedly presenting the unknown stimulus (e.g., math fact) in between the presentation of known stimuli. Higher OTR was shown to result in greater skill acquisition than lower OTR (Sadokierski & Burns, 2008; Zaslofsky et al., 2016). Second, the interspersal of known and unknown stimuli incorporates spacing, which has been shown to support long-term skill maintenance (Rohrer & Taylor, 2007; Swehla et al., 2016). Third, in IR, students typically must generate responses to stimuli, which has been shown to increase skill acquisition (Zaslofsky et al., 2016).

| Step | Fact | Step | Fact | Step | Fact | Step | Fact |
|------|------|------|------|------|------|------|------|
| 1    | UF 1 | 12   | KF 2 | 23   | KF 2 | 34   | KF 6 |
| 2    | KF 1 | 13   | KF 3 | 24   | KF 3 | 35   | KF 7 |
| 3    | UF 1 | 14   | KF 4 | 25   | KF 4 | 36   | UF 1 |
| 4    | KF 1 | 15   | UF 1 | 26   | KF 5 | 37   | KF 1 |
| 5    | KF 2 | 16   | KF 1 | 27   | KF 6 | 38   | KF 2 |
| 6    | UF 1 | 17   | KF 2 | 28   | UF 1 | 39   | KF 3 |
| 7    | KF 1 | 18   | KF 3 | 29   | KF 1 | 40   | KF 4 |
| 8    | KF 2 | 19   | KF 4 | 30   | KF 2 | 41   | KF 5 |
| 9    | KF 3 | 20   | KF 5 | 31   | KF 3 | 42   | KF 6 |
| 10   | UF 1 | 21   | UF 1 | 32   | KF 4 | 43   | KF 7 |
| 11   | KF 1 | 22   | KF 1 | 33   | KF 5 | 44   | KF 8 |

Fig. 1 Incremental rehearsal (IR) sequence. a shuffle incremental rehearsal (ShIR) followed the same sequence as IR above except known facts were shuffled in between each sequence, UF = unknown fact, KF = known fact.
Understanding the theoretical mechanisms underlying interventions is imperative to support intervention effectiveness (Mercer et al., 2014; Varma & Schleisman, 2014) and to guide intervention modifications that do not undermine intervention effectiveness and may actually maximize effectiveness (Burns, 2011). Prior research has examined modified IR approaches by altering the number of OTRs (Volpe et al., 2011; Zaslofsky et al., 2016), response generation (Zaslofsky et al., 2016), stimuli spacing (Swehla et al., 2016), and type of stimulus (Adams & Maki, 2021). Zaslofsky et al. (2016) found that a high OTR condition led to greater retention than a low OTR condition and generating a stimulus response resulted in greater retention than the correct response being included on the stimulus. Swehla and colleagues (2016) found that traditional IR spacing was more effective than exponentially increasing spacing (i.e., one unknown, one known, one unknown, two known, one unknown, four known, one unknown, eight known items). However, providing a visual representation on stimuli (i.e., on a math fact flashcard) did not increase multiplication fact retention (Adams & Maki, 2021), perhaps due to intervention procedural modifications not in alignment with IR’s underlying causal mechanisms. Taken together, this research suggested that manipulating and modifying IR components is a useful way to understand its underlying causal mechanisms and is a promising approach to maximize intervention effectiveness.

As noted above, IR requires students to repeatedly respond to stimuli, and this repeated practice has been shown to be an important intervention component underlying its effectiveness (Szdokierski & Burns, 2008; Zaslofsky et al., 2016). However, because known and unknown items are presented in the same order with each IR sequence, it is possible that students do not truly rehearse items, but rather simply memorize the correct response order. Item rehearsal has been shown to strengthen the relationship between the stimulus and the response (McNamara & Healy, 1995; Rittle-Johnson & Kmicikewycz, 2008). Thus, simply memorizing a response order rather than rehearsing the item may negatively impact learning, with repeated fact rehearsal potentially representing an underlying causal mechanism of fact retention and maintenance. A weak stimulus–response relationship could negatively affect long-term skill maintenance. Thus, research examining IR procedures accounting for potential response order memorization is needed. Further, although there is a relatively strong research-base supporting the effectiveness of IR, relatively little is known about its longer-term effects because most studies have only assessed short-term retention or fluency, underscoring the need for research examining maintenance effects of IR.

**Virtual Intervention Implementation**

The prevalence of virtual technologies to provide support services has been rapidly growing across professional disciplines, including education, with the COVID-19 pandemic promptly propelling the use of virtual platforms to support service provision throughout the USA and worldwide. Even prior to the pandemic, 212,311 students chose to attend fully virtual or online schools (NCES, 2019), with approximately 18.66 billion dollars allotted by the USA in 2019 to global EdTech
investments (Li & Lalani, 2020). However, most research examining virtual learning has involved behavioral health (e.g., anti-bullying; Mayworm et al., 2020), with limited research examining instruction in academic skills in virtual formats. Choi and Walters (2018) evaluated the utility of synchronous small group discourse sessions pertaining to online math learning with elementary students. Students who participated in more virtual sessions were more likely to score at the “Proficient” and “Advanced” levels on standardized state math assessments and earn significantly higher final math course grades at the end of the year (Choi & Walters, 2018), suggesting that virtual math support for elementary students can be beneficial. However, there is limited research examining academic interventions implemented virtually with students who need additional support.

As the use of synchronous and asynchronous virtual technology continues to accelerate within K-12 education, the need for evidence-based virtual intervention procedures also continues to intensify. The evolution of the use of telehealth technologies within schools after the COVID-19 pandemic is not yet known; however, there is reason to believe that a large number of students will opt into attending online or virtual schools (Darling-Hammond et al., 2020). Thus, understanding how instruction and intervention can be effectively provided via technology to students in need of academic supports is imperative and represents one of the next frontiers in intervention research (Goddard et al., 2021). Examining implementation of virtually delivered academic interventions is necessary to ensure that all students receive appropriate academic support services in school.

**Purpose**

There are many students who do not demonstrate grade-level proficiency in math (NCES, 2019), and many students with MD struggle with math fact acquisition and fluency (Calhoon et al., 2007; Nelson et al., 2022), underscoring the need to intervene with math facts to support children’s overall math proficiency (McVancel et al., 2018). Although prior research has supported the use of IR to improve students’ math fact retention, IR procedures may result in students simply memorizing response order, rather than actually rehearsing the math facts, which could negatively impact long-term skill maintenance due to a weak stimulus–response relationship. Further, as the use of virtual technologies in education is likely to continue to increase in schools, examining virtual intervention implementation is needed. The purpose of this study was to examine the differential effectiveness of virtually implemented traditional IR and a modified IR in which known multiplication facts were shuffled between each fact sequence (shuffle IR; ShIR) to investigate potential underlying causal mechanisms. The following research questions guided the study: (1) What is the comparative effect of traditional IR and ShIR on multiplication fact retention, (2) What is the comparative effect of traditional IR and ShIR on multiplication fact maintenance, (3) To what extent are there differences in cumulative errors made during intervention sessions between IR and ShIR intervention conditions?
Method

Participants and Setting

This study was conducted in partnership with a university laboratory school located in a Southeastern mid-size county in the USA. All interventions were implemented virtually, and during the period the study was conducted, students received all instruction virtually due to the COVID-19 pandemic. For the 2019–2020 school year, the school served 1,140 students from kindergarten through 12th grade. Of the 1,140 students, 51% \((n = 583)\) were boys and 49% \((n = 557)\) were girls. Approximately 43.42% \((n = 495)\) of students identified as White, 25.26% \((n = 288)\) identified as Black, 20.08% \((n = 229)\) identified as Hispanic, 7.28% \((n = 83)\) identified as two or more races, 3.3% \((n = 38)\) identified as Asian, 0.52% \((n = 6)\) identified as American Indian/Alaskan Native, and 0.09% \((n = 1)\) identified as Native Hawaiian/Pacific Islander. Thirty-five percent \((n = 394)\) of students were eligible for free/reduced-price lunch (National Center for Education Statistics [NCES], National Center for Education Statistics, 2021).

Students were eligible for potential participation in the study based on two criteria. First, students needed to demonstrate a need for a multiplication fact intervention based on below grade-level benchmark math scores on school-developed benchmark assessments and teacher referral. The school’s math benchmark assessment included items reflecting the grade-level state math standards that students were expected to master across the school year. Because the school math benchmark assessment was a broad measure of math skills, teachers were also asked to refer students who were specifically struggling to acquire multiplication facts. Consent forms were distributed to the parents of ten fourth \((n = 4)\) and fifth grade \((n = 6)\) students meeting those criteria. Parental consent was obtained and parents provided participant demographic information for six students. Second, to participate, students needed to demonstrate at least 72 (out of 144) unknown facts to ensure an adequate number of multiplication facts to rehearse and assess (control condition) during the study. All six students who returned consent forms met this criterion.

Participants included two fourth-grade and four fifth-grade students. The names used to identify each participant are pseudonyms. The first participant, Cami, identified as a White girl in the fifth grade. At baseline, Cami correctly identified 59 multiplication facts. The second participant, Ciera, identified as a Black girl in the fourth grade. Ciera correctly responded to 50 multiplication facts at baseline. The third participant, Carly, identified as a White girl in the fifth grade. Carly correctly identified 54 multiplication facts at baseline. The fourth participant, Indigo, identified as a White and Hispanic boy in the fifth grade. Indigo correctly responded to 54 of 144 multiplication facts at baseline. The fifth participant, Lucy, identified as a White girl in the fourth grade. At baseline, Lucy correctly identified 58 of 144 multiplication facts. Lastly, the sixth participant, Macy, identified as a White and Hispanic girl in the fifth grade. At baseline, Macy correctly identified 51 of 144 multiplication facts.
Measures

Each participant’s multiplication fact skills were assessed at baseline to determine their known and unknown multiplication facts, and multiplication fact skills were frequently assessed over the course of the study to measure multiplication fact retention and maintenance.

Baseline Known and Unknown Facts

Participating students were assessed on 144 multiplication facts (facts 1 through 12) at baseline to determine students’ known and unknown multiplication facts, following Burns’ (2005) procedures. Both commutative property forms of the multiplication facts were assessed (e.g., $3 \times 4$ and $4 \times 3$). The participants were presented one multiplication fact at a time in a random order via PowerPoint, with one fact presented vertically in 72-point font on each slide. The PowerPoint slides were shared via Zoom (Zoom Video Communications, San Jose, CA) to each participant in full screen mode and were the only program on the student’s screen. A response was considered correct if the participant responded correctly within 3 s of the fact being presented. An incorrect response was recorded if the participant provided an incorrect response, responded after the 3 s mark, or did not give a response at all. No corrective feedback was provided.

Retention

Apart from the first intervention session, multiplication fact retention was assessed at the start of each intervention session to measure participants’ retention of multiplication facts that were rehearsed during the previous intervention session. The retention assessment was conducted using a PowerPoint presentation with one multiplication fact presented vertically in 72-point font on each slide following the same procedures as the baseline assessments. A multiplication fact was defined as retained if the participant provided a correct response within 3 s of the multiplication fact’s presentation.

Maintenance

To assess multiplication fact maintenance, the participants were assessed on their multiplication fact recall at one week, two weeks, and three weeks after the completion of the last intervention session by the first author. For example, if a participant’s last intervention session was on Thursday April 22nd, then their maintenance assessments took place on Thursday April 29th, Thursday May 6th, and Thursday May 13th. All students’ rehearsed multiplication facts were assessed during the maintenance sessions. In line with the baseline and retention assessment procedures, multiplication facts were recorded as correct if the student
responded correctly within 3 s, and incorrect if an incorrect, delayed (after 3 s), or no response was provided.

**Study Conditions**

Each student participated in two intervention conditions, IR and ShIR, and a control condition. Participants were randomly assigned to the order in which they completed the conditions. Three students participated in IR first and ShIR second (Cami, Ciera, Indigo) while the other three students participated in ShIR first and IR second (Carly, Lucy, Macy). The control condition was implemented in between the two intervention conditions. Students received each condition four days per week for two weeks each, totaling six weeks across all conditions. Both the IR and ShIR interventions were implemented following a scripted intervention protocol. Intervention implementation fidelity and interobserver agreement were monitored to ensure the intervention was appropriately implemented and student responses were correctly scored as correct or incorrect. For both intervention conditions, each multiplication fact was presented vertically, in full screen, on PowerPoint slides through custom slide shows, allowing for automatic interspersal of participants’ known and unknown multiplication facts.

**Incremental Rehearsal**

The interventionists followed the scripted intervention protocol and first taught the unknown multiplication fact by stating “BLANK times BLANK equals BLANK. What does BLANK times BLANK equal?” If the participant responded correctly in under 3 s, the interventionist responded, “That’s right!” If the student responded with the incorrect answer, the interventionist provided corrective feedback, saying, “That’s not quite right. Let’s try again. BLANK times BLANK equals BLANK. What does BLANK times BLANK equal?” Next, the first known fact was presented to the student (but not taught) and standardized corrective feedback was provided when needed. The intervention sequence was as follows: unknown fact 1, known fact 1, unknown fact 1, known fact 1, known fact 2, unknown fact 1, known fact 1, known fact 2, known fact 3, etc. (see Fig. 1 for complete sequence). This sequence continued until eight known facts were presented. For sequence 2, unknown fact 2 was presented, unknown fact 1 became known fact 1, and the last known fact was dropped from the sequence (Tucker, 1989). Students were taught four unknown facts each day resulting in four IR sequences.

**Shuffle Incremental Rehearsal**

ShIR followed the same general procedures as IR except that participants’ known facts were randomly shuffled between each intervention sequence. Thus, the first unknown fact was taught following the exact same procedures as traditional IR. After the first unknown fact was taught, the first unknown fact became the first known fact and unknown fact 2 entered the sequence. All original known facts were
then shuffled in the PowerPoint so that they were presented in a different order than during the first unknown fact sequence. Unknown facts were not shuffled to maintain the same OTRs between the IR and ShIR intervention conditions. Like IR, a scripted intervention protocol was followed for the ShIR condition using the same language as the IR condition, except all known multiplication facts were shuffled between sequence.

Control Condition

In the control condition, participants were assessed on four unique multiplication facts during each session following the same procedures used to identify known and unknown facts at baseline. The interventionist did not model or teach the multiplication facts. The interventionist presented each fact one time to the student via individual PowerPoint slides. If the student correctly responded to the multiplication fact within 3 s, the multiplication fact was marked as correct. If the participant did not correctly respond to the multiplication fact within 3 s, the multiplication fact was marked as incorrect. Participants were not told whether they responded correctly or incorrectly to the fact.

Procedures

University Institutional Review Board approval, parental consent, and child assent were obtained prior to all data collection. Baseline data were collected by individually assessing each participants’ knowledge of 144 multiplication facts (facts 1 through 12) and recording the fact as either known or unknown as described above. Each participant’s unknown multiplication facts were compiled and randomly assigned to be rehearsed in the IR or ShIR intervention conditions or to the control condition. The control condition facts were not rehearsed, but were assessed over the course of the study as a comparison to the cumulative retention in the IR and ShIR conditions. Each condition’s multiplication facts were randomly selected to be taught for each intervention session (or assessed in the control condition) using a random number generator. Students were randomly assigned to the order in which they participated in the IR or ShIR conditions prior to the start of intervention sessions, with three participants receiving IR first and three participants receiving ShIR first. Participants completed eight sessions in each condition, with retention assessed eight times for IR, ShIR, and control conditions.

All intervention sessions were implemented virtually via Zoom during participants’ regularly scheduled intervention time, writing class, specials (i.e., physical education, art, music), or other time frames in which the students were not receiving core math or reading instruction. Participating students completed each study session from a quiet space in their home (e.g., desk in bedroom). Each intervention session meeting began by virtually moving students into a private focus Zoom session with an interventionist and fidelity assessor (for 25% of sessions) via Hāpara through Google Classroom (Google, Mountain View, CA). The focus session capability provided by Hāpara allowed the interventionist to limit each student from viewing or
interacting with any other programs during all intervention sessions. Each IR and ShIR intervention session lasted an average of 21.51 min, while control condition sessions lasted an average of 5.68 min.

During the interventionist’s first session with a participant, they spent several minutes developing rapport with the student by talking about school and student interests. Then, in all intervention sessions except the first session, the interventionist used the Zoom screen share function to present a PowerPoint containing the four multiplication facts rehearsed during the prior intervention session to assess their retention of the previous session’s rehearsed multiplication facts. The interventionist stated, “I am going to see if you remember the multiplication facts we worked on (last day intervention was implemented).” If the student correctly responded to the multiplication fact within 3 s, the multiplication fact was marked retained and not rehearsed again. If the participant did not correctly respond to the multiplication fact within 3 s, the multiplication fact was not considered retained and the fact was rehearsed again in the subsequent intervention session.

Participants rehearsed four unknown multiplication facts during each IR and ShIR intervention session, which included new multiplication facts or previously rehearsed multiplication facts if not retained during the retention assessment. The interventionist screen shared the participant’s individualized IR or ShIR intervention PowerPoint according to their current intervention condition via Zoom. The interventionist started a timer prior to introducing the first multiplication fact and stopped the timer after the last multiplication fact was rehearsed to account for the total time of the IR and ShIR intervention sequences. The intervention was then implemented following the scripted procedures. After the IR or ShIR intervention was completed, the session timer was stopped, and the corresponding time for the intervention was recorded.

Implementation Fidelity and Interobserver Agreement

Graduate students in school psychology and the first author, who holds a Ph.D. in Educational Psychology, implemented the interventions. All interventionists implemented both the IR and ShIR interventions. The first author led a two-hour training session on all study procedures with the interventionists prior to the start of the study. Interventionists rehearsed intervention implementation until demonstrating 100% implementation fidelity and correct multiplication fact response scoring. The first author and graduate student interventionists observed 25% of all intervention sessions (two intervention sessions per intervention condition per participant). Implementation fidelity was calculated by dividing the total number of intervention steps that were correctly implemented by the total number of intervention steps (correctly implemented plus incorrectly implemented intervention steps). Implementation fidelity ranged from 92.31 to 100% and averaged 98.42% across all conditions, participants, and interventionists. Interobserver agreement (IOA) was also calculated for scoring of the dependent variables, multiplication fact retention and maintenance, by dividing the number of scoring agreements by scoring agreements.
plus scoring disagreements. IOA across conditions, participants, and interventionists ranged from 88.45 to 100%, averaging 96.24%.

Research Design and Data Analyses

The research questions were addressed using a single-case cumulative acquisition design, which is a repeated acquisition design that can be used to determine the relative effectiveness and efficiency of two or more interventions (Riley-Tillman et al., 2020). Like repeated acquisition designs, cumulative acquisition designs are useful for examining the effects of interventions on nonreversible behaviors, such as multiplication fact skills (Kirby et al., 2021). Although repeated and cumulative acquisition designs have been criticized due to threats to internal validity given no baseline condition is included, randomizing stimuli to conditions, counterbalancing treatment order, and including control stimuli support internal validity (Kirby et al., 2021), all of which were included in this study. Participants’ cumulative retention and maintenance of multiplication facts and cumulative errors made during intervention sessions were graphed and are presented in Figs. 2, 3, 4 and 5, respectively. We used visual analysis to examine level, trend, and overlap of the data, and we fit generalized linear models (GLM) to the data using generalized least squares (GLS) to calculate slope for cumulative multiplication fact retention and cumulative errors for each participant. Given the single-case design used in this study, GLS was used to compute slope to account for autocorrelation of the data (Somer et al., 2022). The models we fit can be noted as:

\[ Y_t = b_0Y + b_{1\text{level}} \text{phase}_t + b_{1\text{trend}} \text{time}_t + b_{2\text{level}} \text{phase}_t + b_{2\text{trend}} \text{time}_t + b_{3\text{level}} \text{phase}_t + b_{3\text{trend}} \text{time}_t + e_{Y,t} \]

where, \( Y_t \) = cumulative multiplication fact retention and cumulative errors, \( b_0Y \) = the intercept or initial data point in the phase (IR, ShIR, or Control), \( b_{1\text{level}} \) = level in first phase (IR, ShIR, or Control), \( b_{1\text{trend}} \) = trend in first phase, \( b_{2\text{level}} \) = level in second phase, \( b_{2\text{trend}} \) = trend in second phase, \( b_{3\text{level}} \) = level in third phase, \( b_{3\text{trend}} \) = trend in third phase, \( e_{Y,t} = p_Y e_{Y,t-1} + \nu_{Y,t} \) to account for autocorrelation among the errors (\( p_Y \) = strength of serial dependency and \( \nu_{Y,t} \) = error term).

The GLMs were fit in R using the glm2 package (Pustejovsky & Swan, 2018).

Results

Cumulative Multiplication Fact Retention

All participants demonstrated an increase in level and an increasing trend in multiplication facts retained in both the IR and ShIR conditions, and IR and ShIR consistently resulted in greater level and trend in multiplication fact retention than the control condition. However, differences in level and trend between the IR and ShIR conditions were not consistent across participants.
Over the course of the interventions, Cami’s level of cumulative retained facts was consistently higher in the ShIR condition (25 facts retained at session 8) than in the IR condition (22 facts retained at session 8), with no overlap among conditions. Cami did not accurately respond to any multiplication facts in the control condition over the course of the study. The trend of Cami’s cumulative retained facts was the same in the IR and ShIR conditions until session 6, when greater trend was evident in the ShIR condition compared to the IR condition. Cami’s slope in the ShIR condition was slightly higher (3.14 facts gained per session) compared to the IR condition (2.82 facts gained per session).

Fig. 2  Cumulative multiplication fact retention for Cami, Ciera, and Carly, Main = maintenance; Indigo intervention 1 = IR, intervention 2 = ShIR; Lucy intervention 1 = ShIR, intervention 2 = IR; Macy intervention 1 = ShIR, intervention 2 = IR
Ciera’s data showed greater level and trend in the IR condition compared to the ShIR and control conditions, with some initial overlap between the two intervention conditions until session 3. Ciera’s cumulative retention in both the IR and ShIR conditions also showed greater level and trend compared to the control condition. In the IR condition, Ciera retained 21 multiplication facts at session 8, but she retained 12 multiplication facts in the ShIR condition and 3 multiplication facts in the control condition at session 8. The trend of Ciera’s data was also greatest in the IR condition with a slope of 2.80, followed by the ShIR condition with a slope of 1.17, and the control condition with a slope of 0.50.
The level and trend of Carly’s data were the same in the IR and ShIR conditions until session 5, when the level of her ShIR (25 multiplication facts retained at session 8) data was consistently higher than her IR data (22 multiplication facts retained) through session 8. Both Carly’s IR and ShIR data showed greater level than in the control condition in which Carly correctly responded to one multiplication fact. The trend of Carly’s data was also the same in the IR and ShIR conditions until session 5 when the trend increased in the ShIR condition compared to the IR condition. The slope of Carly’s data was 3.60 in the ShIR condition, 2.96 in the IR condition, and 0.18 in the control condition. Over the course of the interventions, the level of Indigo’s IR data (22 multiplication facts retained at session 8) was
consistently higher than the level of his ShIR data (18 multiplication facts retained at session 8) and control condition data (3 multiplication facts at session 8), with no overlap among conditions. The trend of Indigo’s data was similar in both the IR and ShIR conditions until session 6 when his trend in the IR condition increased compared to the ShIR condition. Trend in both the IR (slope of 2.64) and ShIR (slope of 2.40) conditions was greater than the control (slope of 0.48) condition.

The level of Lucy’s cumulative retained facts was consistently greater in the IR condition (27 multiplication facts retained in session 8) compared to the ShIR condition (20 multiplication facts retained at session 8), which was consistently greater than the control condition (2 multiplication facts at session 8), and there was no overlap among conditions. The trend of Lucy’s data was similar in both the IR and ShIR conditions until session 6 when his trend in the IR condition increased compared to the ShIR condition. Trend in both the IR (slope of 2.64) and ShIR (slope of 2.40) conditions was greater than the control (slope of 0.48) condition.

The level of Macy’s cumulative retained facts was consistently greater in the IR condition (55 multiplication facts retained in session 8) compared to the ShIR condition (45 multiplication facts retained at session 8), which was consistently greater than the control condition (5 multiplication facts at session 8), and there was no overlap among conditions. The trend of Macy’s data was similar in both the IR and ShIR conditions until session 6 when his trend in the IR condition increased compared to the ShIR condition. Trend in both the IR (slope of 2.64) and ShIR (slope of 2.40) conditions was greater than the control (slope of 0.48) condition.

**Fig. 5** Cumulative intervention session errors for Indigo, Lucy, and Macy
overlap among conditions. The trend of Lucy’s cumulative retained facts was greater in the IR and ShIR conditions compared to the control condition (slope of 0.37), and the trend in the IR and ShIR conditions was similar until session 6 when the trend in the ShIR condition (slope of 2.70) was less than the trend in the IR condition (slope of 3.32). The level of Macy’s data was greater in the ShIR condition (27 multiplication facts retained at session 8) compared to the IR (18 multiplication facts retained at session 8) and control conditions (2 multiplication facts at session 8). The trend of Macy’s data was also greatest in the ShIR condition (slope of 3.45) compared to the IR condition (slope of 2.07) and control condition (slope of 0.19).

**Multiplication Fact Maintenance**

Regardless of which intervention condition resulted in the greatest cumulative multiplication fact retention, participants generally demonstrated greater multiplication fact maintenance in the ShIR condition. Cami maintained all 25 rehearsed multiplication facts in the ShIR condition across all three maintenance assessments, and she maintained 22 multiplication facts at the first maintenance assessment and 20 multiplication facts (out of 22 total rehearsed) at the second and third maintenance assessments, respectively, in the IR condition. Ciera maintained all 12 multiplication facts rehearsed in the ShIR condition for each of the three maintenance assessments. She maintained 20 out of 20 multiplication facts at the first maintenance assessment and 18 (out of 20) multiplication facts at the second and third maintenance assessments in the IR condition. Carly maintained 100% of rehearsed multiplication facts in the ShIR (29 facts) and IR (25 facts) conditions across all three maintenance assessments.

In the ShIR condition, Indigo maintained 17 (out of 17) rehearsed multiplication facts at maintenance assessment one and 16 multiplication facts at maintenance assessments two and three, respectively. He maintained 22 (out of 22) rehearsed multiplication facts at maintenance assessment one, 20 at maintenance assessment two, and 18 at maintenance assessment three in the IR condition. Lucy maintained 20, 19, and 19 multiplication facts at the first, second, and third maintenance assessments, respectively, in the ShIR condition, and 27, 26, and 25 multiplication facts at the first, second, and third maintenance assessments, respectively, in the IR condition. Macy maintained 27 (out of 27) multiplication facts at all three maintenance assessments in the ShIR condition and 18 multiplication facts (out of 18) at the first maintenance assessment and 15 multiplication facts at the second and third maintenance assessments in the IR condition.

**Cumulative Errors**

Differences in level and trend of cumulative error data between the IR and ShIR conditions were also inconsistent across participants. However, all participants demonstrated an inverse in cumulative errors from cumulative multiplication facts retained. That is, participants whose cumulative retention data showed greater level
and trend in the IR condition demonstrated greater level and trend in the ShIR condition for cumulative intervention session errors and vice versa.

Beginning in session 3 and continuing through session 8, Cami’s cumulative error data in the ShIR condition demonstrated greater level and trend than in the IR condition. By session 8, Cami made 41 errors during the IR intervention sessions and 32 errors in the ShIR intervention sessions. The slope of her IR data was 5.33, and the slope of her ShIR data was 3.94. The level (IR session 8 errors = 19, ShIR session 8 errors = 36) and trend (IR slope = 2.55, ShIR slope = 5.39) of Ciera’s data were greater in the ShIR condition compared to the IR condition, with overlap between the two conditions until session 3. The level of Carly’s cumulative error data was initially greater in the ShIR condition until session 4, when the level was greater in the IR condition than the ShIR condition through session 8. By session 8, Carly made 34 errors during the IR intervention sessions and 24 errors in the ShIR intervention sessions. The slope of her IR data was 5.15, and the slope of her ShIR data was 3.20.

The level of Indigo’s cumulative error data was similar in both the IR and ShIR conditions over the course of the study, with 45 cumulative errors at session 8 in the IR condition and 43 cumulative errors in the ShIR condition. The trend of his data was slightly higher in the IR condition (slope = 6.00) than in the ShIR condition (slope = 5.33). The level of Lucy’s data was consistently higher in the ShIR condition (51 cumulative errors at session 8) compared to the IR condition (37 cumulative errors at session 8). Lucy’s data trend was also greater in the ShIR condition (slope = 5.87) than the IR condition (slope = 4.25). Macy’s cumulative error data in the ShIR condition demonstrated greater level and trend than in the IR condition. By session 8, Macy made 51 errors during the IR intervention sessions and 40 errors in the ShIR intervention sessions. The slope of her IR data was 7.14, and the slope of her ShIR data was 5.75.

Discussion

This study examined the differential effectiveness of IR and a modified IR, ShIR, in which known multiplication facts were shuffled in between each IR sequence using a cumulative acquisition design with six fourth- and fifth-grade participants with multiplication fact difficulties. Examining modified intervention procedures helps with understanding the intervention’s underlying causal mechanisms and represents a promising approach to maximize intervention effectiveness (Burns, 2011). This study, therefore, examined a modified IR intervention based on underlying theoretical mechanisms. We sought to understand if ShIR could support greater fact retention and maintenance due to greater math fact rehearsal in ShIR compared to response order memorization in IR, which could result in a stronger relationship between the stimulus and response in the ShIR condition and consequently, greater retention and maintenance.

All six participating students demonstrated greater level and trend in the IR and ShIR conditions compared to the control condition. However, consistent differences
in level and trend between the IR and ShIR conditions were not evident across participants, with IR resulting in greater multiplication fact retention for three participants (Ciera, Indigo, and Lucy) and ShIR resulting in greater retention for three participants (Cami, Carly, and Macy). Further, all participants made more cumulative errors during intervention sessions in the intervention condition with lower cumulative fact retention. That is, there was an inverse relationship between cumulative retention and cumulative intervention errors, with participants retaining more multiplication facts in the intervention condition in which they made fewer intervention session errors.

**Multiplication Fact Retention and Maintenance**

Drill-based interventions have been shown to be one of the most effective approaches to target math fact retention and fluency (Burns et al., 2010), and this study’s findings coincided with prior research showing that IR has consistently demonstrated effectiveness for increasing math fact acquisition and fluency (Burns et al., 2010; Codding et al., 2009, 2011). Moreover, studies examining modified IR procedures have led to greater understanding of its underlying causal mechanisms (e.g., Swehla et al., 2016; Zaslofsky et al., 2016). However, studies comparing the effectiveness of IR with modified IR approaches have not consistently shown IR or modified IR approaches to be more effective than the other (Adams & Maki, 2021; Swehla et al., 2016; Volpe et al., 2011). Some prior research examining IR modifications has altered IR procedures in ways that were inconsistent with the underlying causal mechanisms. For example, some modifications have resulted in fewer OTR to stimuli, which is critically important for acquisition and retention of math facts and other basic skills (Adams & Maki, 2021; Swehla et al., 2016; Volpe et al., 2011; Zaslofsky et al., 2016). Consequently, in this study, OTRs were held constant between the traditional IR and ShIR intervention conditions, but similar findings were evident in this study as some participants made greater gains in response to the IR intervention condition while other participants made greater gains in the ShIR condition. Finally, students generally did not respond correctly to multiplication facts in the control condition, suggesting that the explicit instruction and rehearsal incorporated in the intervention conditions was necessary to promote retention and maintenance rather than simply testing of math facts. However, this hypothesis should be further examined with a larger sample.

Modified intervention approaches represent a promising approach to maximize intervention effectiveness. With IR, students are repeatedly shown the same stimuli in the same order, possibly leading some students to simply memorize the correct answer rather than responding to the stimulus, which may undermine students’ rehearsal and long-term maintenance of the information (Rittle-Johnson & Kmicikewycz, 2008). In this study, therefore, we examined a modified IR procedure in which known multiplication facts were shuffled in between each fact sequence. However, the mixed findings in this study in which half of the participants retained more multiplication facts in the IR condition while the other half of participants retained more facts in the ShIR condition suggested that this modified approach was
not necessarily more effective than IR for short-term retention. For some students, it may be that shuffling known facts actually increased taxing of working memory and cognitive load, thereby interfering with students’ acquisition and retention of the multiplication facts. Further, with IR, it may be that interspersal of known and unknown information supports fact retention without straining cognitive resources (Swehla et al., 2016; Varma & Schleisman, 2014) regardless of the specific stimulus or stimulus order. Interestingly, though, most participants demonstrated more stable multiplication fact maintenance in the ShIR condition than in the IR condition, regardless of the intervention condition that resulted in greater cumulative fact retention. Thus, multiplication fact rehearsal within ShIR (compared to multiplication fact response order in IR) may support longer-term skill maintenance. Further research is needed to explore these hypotheses.

**Intervention Errors**

All participants made more cumulative errors during intervention sessions in the intervention condition in which they retained fewer multiplication facts. For example, if a participant exhibited higher cumulative multiplication fact retention in the ShIR condition, they made more cumulative errors during IR intervention sessions. These findings were not consistent with prior research showing that intervention errors were not related to multiplication fact retention (Adams & Maki, 2021). However, in their study, Adams and Maki compared IR, a modified IR approach, and traditional drill (TD) and found that all participants made the most errors in the TD condition; thus, their findings may be specific to that intervention approach and may not be directly comparable to this study’s findings.

Accurate stimulus responding and skill rehearsal is important to support skill retention, maintenance, and fluency. High error rates may lead to rehearsal of inaccurate information, resulting in retention of incorrect responses (Kupzyk et al., 2012). In this study, when students made more errors during intervention sessions, they retained fewer multiplication facts. Student response errors have been shown to increase when rehearsing too much information, thus taxing working memory and other cognitive resources (Burns et al., 2016). In this study, participants rehearsed the same amount of information in both intervention conditions, and it is unclear why some participants made more errors in the IR condition while others made more errors in the ShIR condition. However, it may be that participants exhibited individual differences in which intervention condition resulted in the greater cognitive load, with some students’ cognitive resources being more taxed in the IR condition and others’ cognitive resources being more taxed in the ShIR condition. Additional research employing research designs allowing for examination of such individual differences is warranted.

**Limitations and Implications for Research**

The study findings should be interpreted within the context of the study limitations. First, the interventions were implemented virtually, which may have resulted
in confounding variables (e.g., participant attention, minor technology glitches) affecting the study findings. Although we employed a relatively novel approach to academic intervention implementation, in part necessitated by the COVID-19 pandemic, it is not clear if results would be similar when implemented using traditional IR procedures in a face-to-face setting. Additional research examining these interventions with traditional procedures is therefore warranted. Moreover, because it seems likely that virtual schooling will continue beyond the pandemic with increasing numbers of virtual school options (Darling-Hammond et al., 2020), further research examining virtual interventions is needed to support students with academic needs attending such schools.

Second, we included the two commutative property forms of all multiplication facts in this study. To examine the differential effectiveness of the IR and ShIR intervention conditions, we randomly assigned participants’ unknown multiplication facts to the IR, ShIR, and control conditions. But these procedures may have supported participants’ learning of the other commutative property form in a subsequent intervention session if one form had previously been rehearsed. Excluding one commutative property form of multiplication facts could be examined in future research. Third, we were only able to examine maintenance up to three weeks following interventions due to the end of the school year. Thus, it is not known if participants’ maintenance of the multiplication facts continued over an even longer period, which, of course, is the ultimate goal of such interventions.

Fourth, all students were taught four multiplication facts in each intervention session for both the IR and ShIR intervention conditions. However, we did not assess participants’ acquisition rates, or the amount of information they could learn in a session (Burns et al., 2016). Average acquisition rates range from approximately 4.5–6.6 facts per intervention session for students in fourth and fifth grades (Haegele & Burns, 2015). In this study, we rehearsed four new multiplication facts because all students had multiplication difficulties. However, for some participants, four multiplication facts may have been too many to rehearse in one intervention session, which could have led to a high error rate, impacting retention (Haegele & Burns, 2015). If four multiplication facts were too few to rehearse in one session, then intervention effectiveness may have been undermined because students could have learned more multiplication facts than they had the opportunity to. Future research could examine these interventions while matching the number of facts taught to participants’ acquisition rates.

Conclusion

This study’s findings highlighted the utility of examining traditional and modified drill interventions in schools. Participants demonstrated greater cumulative multiplication retention level and trend in the IR and ShIR intervention conditions than in the control condition, but half of the participants showed greater retention in the IR condition and half showed greater retention in the ShIR condition, underscoring the need to examine individual student needs when implementing math interventions. Further, this study highlighted the need to further examine IR’s effects on
skill maintenance because the modified IR approach, ShIR, appeared to result in more stable multiplication fact maintenance. Finally, results showed the importance of ensuring low error responding during intervention sessions because participants retained more multiplication facts in the intervention condition in which they made fewer intervention session errors. Research examining intervention modifications furthers understanding of causal mechanisms underlying intervention effectiveness, which could lead to increased intervention effectiveness.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures were approved by and in accordance with the ethical standards of the University Institutional Review Board and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

Adams, S. R., & Maki, K. E. (2021). Examining the differential effectiveness and efficiency of alternative multiplication drill interventions with third-grade students. *Journal of Applied School Psychology, 37*, 352–376. https://doi.org/10.1080/15377903.2020.1848956

Ardoin, S. P., Binder, K. S., Zawoyski, A. M., & Foster, T. E. (2018). Examining the maintenance and generalization effects of repeated practice: A comparison of three interventions. *Journal of School Psychology, 68*, 1–18.

Ardoin, S. P., & Daly, E. J. (2007). Introduction to the special series: Close encounters of the instructional kind—how the instructional hierarchy is shaping instructional research 30 years later. *Journal of Behavioral Education, 16*, 1–6.

Bailey, D. H., Siegler, R., & Geary, D. (2014). Early predictors of middle school fraction knowledge. *Developmental Science, 17*, 775–785. https://doi.org/10.1111/desc.12155

Burns, M. K. (2005). Using incremental rehearsal to increase fluency of single-digit multiplication facts with children identified as learning disabled in mathematics computation. *Education and Treatment of Children, 28*, 237–249.

Burns, M. K. (2011). School psychology research: Combining ecological theory and prevention science. *School Psychology Review, 40*(1), 132–140. https://doi.org/10.1080/02796015.2011.12087732

Burns, M. K., Codding, R. S., Boice, C. H., & Lukito, G. (2010). Meta-analysis of acquisition and fluency math interventions with instructional and frustration level skills: Evidence for a skill-by-treatment interaction. *School Psychology Review, 39*(1), 69–83. https://doi.org/10.1080/02796015.2010.12087791

Burns, M. K., Zaslowski, A. F., Maki, K. E., & Kwong, E. (2016). Effect of modifying intervention set size with acquisition rate data while practicing single-digit multiplication facts. *Assessment for Effective Intervention, 41*(3), 131–140. https://doi.org/10.1177/1534508415593529

Burns, M. K., Wallick, C., Simonson, G. R., Dominguez, L., Harelstad, L., Kincaid, A., & Nelson, G. S. (2015). Using a conceptual understanding and procedural fluency heuristic to target math interventions with students in early elementary. *Learning Disabilities Research & Practice, 30*(2), 52–60. https://doi.org/10.1111/ldr.12056

Calhoon, M. B., Emerson, R. W., Flores, M., & Houchins, D. E. (2007). Computational fluency performance profile of high school students with mathematics disabilities. *Remedial and Special Education, 28*, 292–303. https://doi.org/10.1177/07419325070280050401
Cameron, C. E., Kim, H., Duncan, R. J., Becker, D. R., & McClelland, M. M. (2019). Bidirectional and co-developing associations of cognitive, mathematics, and literacy skills during kindergarten. *Journal of Applied Developmental Psychology, 62*, 135–144. https://doi.org/10.1016/j.appdev.2019.02.004

Choi, J., & Walters, A. (2018). Exploring the impact of small-group synchronous discourse sessions in online math learning. *Online Learning, 22*(4), 47–64. https://doi.org/10.24059/olj.v22i4.1511

Codding, R. S., Burns, M. K., & Lukito, G. (2011). Meta-analysis of mathematic basic-fact fluency interventions: A component analysis. *Learning Disabilities Research & Practice, 26*(1), 36–47. https://doi.org/10.1111/j.1540-5826.2010.00323.x

Codding, R. S., Hilt-Panahon, A., Panahon, C. J., & Benson, J. L. (2009). Addressing mathematics computation problems: A review of simple and moderate intensity interventions. *Education and Treatment of Children, 32*, 279–312.

Darling-Hammond, L., Schachner, A., Edgerton, A. K. with Badrinarayan, A., Cardichon, J., Cookson, P. W., Jr., Griffith, M., Klevan, S., Maier, A., Martinez, M., Melnick, H., Truong, N., Wojcikiewicz, S. (2020). *Restarting and reinventing school: Learning in the time of COVID and beyond*. Learning Policy Institute. https://edtechbooks.org/-BIT

Delazer, M., Ischebeck, A., Domahs, F., Zamarian, L., Koppelstaetter, F., Siedentopf, C. M., Kaufmann, L., Benke, T., & Felber, S. (2005). Learning by strategies and learning by drill-evidence from an fMRI study. *NeuroImage, 25*(3), 838–849.

Duncan, G., Dowsett, C., Claessens, A., Magnuson, K., Huston, A., Klebanov, P., Pagani, L., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*, 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Fan, W., & Wolters, C. A. (2014). School motivation and high school dropout: The mediating role of educational expectation. *British Journal of Educational Psychology, 84*(1), 22–39. https://doi.org/10.1111/bjep.12002

Goddard, A., Sullivan, E., Fields, P., & Mackey, S. (2021). The future of telehealth in school-based health centers: Lessons from COVID-19. *Journal of Pediatric Health Care, 35*(3), 304–309. https://doi.org/10.1016/j.jpedhc.2020.11.008

Haegele, K., & Burns, M. K. (2015). Effect of modifying intervention set size with acquisition rate data among students identified with a learning disability. *Journal of Behavioral Education, 24*(1), 33–50. https://doi.org/10.1007/s10864-014-9201-0

Haring, N. G., & Eaton, M. D. (1978). Systematic instructional technology: An instructional hierarchy. In N. G. Haring, T. C. Lovitt, M. D. Eaton, & C. L. Hansen (Eds.), *The fourth R: Research in the classroom* (pp. 23–40). Merrill.

Kirby, M. S., Spencer, T. D., & Ferron, J. (2021). How to be RAD: Repeated acquisition design features that enhance internal and external validity. *Perspectives on Behavior Science, 44*, 389–416.

Kupzyk, S., Daly, E. J., III., Ihlo, T., & Young, N. D. (2012). Modifying instruction within tiers in multiltiered intervention programs. *Psychology in the Schools, 49*(3), 219–230.

Lee, J. (2012). College for all: Gaps between desirable and actual P-12 math achievement trajectories for college readiness. *Educational Researcher, 41*, 43–55.

Li, C., & Lalani, F. (2020). *The COVID-19 pandemic has changed education forever: This is how*. World Economic Forum. https://www.weforum.org/agenda/2020/04/coronavirus-education-global-covid19-online-digital-learning

Maki, K. E., Zaslavsky, A. F., Knight, S., Ebbesmeyer, A. M., & Chelmo-Boatman, A. (2021). Intervening with multiplication fact difficulties: Examining the utility of the instructional hierarchy to target interventions. *Journal of Behavioral Education, 30*(4), 534–558. https://doi.org/10.1007/s10864-020-09388-0

Mayworm, A. M., Lever, N., Gloff, N., Cox, J., Willis, K., & Hoover, S. A. (2020). School-based telespsychiatry in an urban setting: Efficiency and satisfaction with care. *Telemedicine Journal and E-Health: The Official Journal of the American Telemedicine Association, 26*(4), 446–454. https://doi.org/10.1089/tmj.2019.0038

McKeveit, N. M., & Codding, R. S. (2021). Brief experimental analysis of math interventions: A synthesis of evidence. *Assessment for Effective Intervention, 46*(3), 217–227. https://doi.org/10.1177/1534508419883937

McNamara, D. S., & Healy, A. F. (1995). A procedural explanation of the generation effect: The use of an operand retrieval strategy for multiplication and addition problems. *Journal of Memory and Language, 34*(3), 399–416. https://doi.org/10.1006/jmla.1995.1018
McVancel, S. M., Missall, K. N., & Bruhn, A. L. (2018). Examining incremental rehearsal: Multiplication fluency with fifth-grade students with math IEP goals. *Contemporary School Psychology, 22*(3), 220–232. https://doi.org/10.1007/s40688-018-0178-x

Mercer, S. H., Idler, A. M., & Bartfai, J. M. (2014). Theory-driven evaluation in school psychology intervention research: 2007–2012. *School Psychology Review, 43*(2), 119–132. https://doi.org/10.1080/02796015.2014.12087439

National Center for Education Statistics (2019). *The nation’s report card: 2019 mathematics results.* Institute for Education Sciences (IES), US Department of Education.

National Center for Education Statistics. (2021). *CCD School Locator.* Retrieved from https://nces.ed.gov/ccd/schoolsearch/

Nelson, G., Crawford, A., Hunt, J., Park, S., Leckie, E., Duarte, A., Bradfoff, T., Ramos-Duke, M., & Zarate, K. (2022). A systematic review of research syntheses on students with mathematics learning disabilities and difficulties. *Learning Disabilities Research & Practice, 37*(1), 18–36. https://doi.org/10.1111/ldr.p.12272

Purpura, D. J., Logan, J. A. R., Hassinger-Das, B., & Napoli, A. R. (2017). Why do early mathematics skills predict later reading? The role of mathematical language. *Developmental Psychology, 53*(9), 1633–1642. https://doi.org/10.1037/dev0000375

Pustejovsky, J. E., & Swan, D. M. (2018). Single-case effect size calculator (Version 0.5) Web application. Retrieved from: https://jepusto.shinyapps.io/SCD-effect-sizes/

Riley-Tillman, T. C., Burns, M. K., & Kilgus, S. P. (2020). *Evaluating educational interventions: Single-case design for measuring response to intervention* (2nd ed.). Guilford.

Rittle-Johnson, B., & Kmicikewycz, A. O. (2008). When generating answers benefits arithmetic skill: The importance of prior knowledge. *Journal of Experimental Child Psychology, 101*(1), 75–81. https://doi.org/10.1016/j.jecp.2008.03.001

Rohrer, D., & Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science, 35*(6), 481–498. https://doi.org/10.1007/s11251-007-9015-8

Somer, E., Gische, C., & Miočević, M. (2022). Methods for modeling autocorrelation and handling missing data in mediation analysis in single case experimental designs (SCEDs). *Evaluation & the Health Professions, 45*, 36–53.

Swehla, S. E., Burns, M. K., Zaslofsky, A. F., Hall, M. S., Varma, S., & Volpe, R. J. (2016). Examining the use of spacing effect to increase the efficiency of incremental rehearsal. *Psychology in the Schools, 53*(4), 404–415. https://doi.org/10.1002/pits.21909

Szadokierski, I., & Burns, M. K. (2008). Analogue evaluation of the effects of opportunities to respond and ratios of known items within drill rehearsal of Esperanto words. *Journal of School Psychology, 46*, 593–609. https://doi.org/10.1016/j.jsp.2008.06.004

Tucker, J. A. (1989). Basic flashcard technique when vocabulary is the goal. [Unpublished teaching materials]. University of Tennessee at Chattanooga.

Vanderheyden, A., & Codding, R. (2020). Belief-based versus evidence-based math assessment and instruction: What school psychologists need to know to improve student outcomes. *NASP Communique, 48*(5), 1.

Varma, S., & Schleisman, K. B. (2014). The cognitive underpinnings of incremental rehearsal. *School Psychology Review, 43*(2), 222–229. https://doi.org/10.1002/spr.21909

Volpe, R. J., Mulé, C. M., Briesch, A. M., Joseph, L. M., & Burns, M. K. (2011). A comparison of two flashcard drill methods targeting word recognition. *Journal of Behavioral Education, 20*(2), 117–137. https://doi.org/10.1007/s10864-011-9124-y

Wang, M. T., & Fredricks, J. A. (2014). The reciprocal links between school engagement, youth problem behaviors, and school dropout during adolescence. *Child Development, 85*(2), 722–737. https://doi.org/10.1111/cdev.12138

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What’s past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher, 43*(7), 352–360. https://doi.org/10.3102/0013189X14553660

Watts, T. W., Duncan, G. J., Chen, M., Claessens, A., Davis-Kean, P. E., Duckworth, K., Engel, M., Siegler, R., & Susperreguy, M. I. (2015). The role of mediators in the development of longitudinal mathematics achievement associations. *Child Development, 86*(6), 1892–1907. https://doi.org/10.1111/cdev.12416

Woodward, J. (2006). Developing automaticity in multiplication facts: Integrating strategy instruction with timed practice drills. *Learning Disability Quarterly, 29*(4), 269–289. https://doi.org/10.12307/30035554
Zaslofsky, A. F., Scholin, S. E., Burns, M. K., & Varma, S. (2016). Comparison of opportunities to respond and generation effect as potential causal mechanisms for incremental rehearsal with multiplication combinations. *Journal of School Psychology, 55*, 71–78.

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