



Does Intervention Session Length Matter? A Comparison of a Math Intervention Across Three Durations

Emily R. DeFouw¹  · Melissa A. Collier-Meek² · Brian Daniels² · Robin S. Coddling³ · Margarida Veiga²

Accepted: 4 November 2021 / Published online: 19 November 2021

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

Abstract

Understanding how to intensify an intervention is critical to implementing response-to-intervention (RtI), a tiered, problem-solving framework to deliver evidence-based interventions per student needs. Research and practice guidelines provide conflicting recommendations for delivering math interventions (i.e., treatment intensity) within RtI. The present study used an adapted alternating treatment design to examine the impact of varying the session length, or number of minutes per intervention session, on student outcomes (i.e., digits correct per min, DCPM). Four female first-grade students identified in the frustrational range (i.e., less than 14 DCPM) for subtraction computation received a packaged intervention (i.e., Cover-Copy-Compare and Schema-Based Word Problem Instruction). The math intervention during the shortest (i.e., 10-min) session length led to similar improvements compared to two of the recommended session length conditions (i.e., 20–40 min) for three out of four students. Following the 5-week intervention, all four students reached an instructional level (i.e., 14–31 DCPM) for the intervention targets, similarly, across all three session length conditions. Implications for research on treatment intensity, math interventions, and RtI are discussed.

Keywords Math intervention · Treatment intensity · Response to intervention

According to the National Center for Education Statistics (NCES, 2018), 60% of 4th grade students and 66% of 8th grade students are considered to be at Basic or Below Basic levels in mathematics, suggesting that these students demonstrate limited

✉ Emily R. DeFouw
Emily.DeFouw@usm.edu

¹ Univeristy of Southern Mississippi, Hattiesburg, MS, USA

² University of Massachusetts Boston, Boston, MA, USA

³ Northeastern University, Boston, MA, USA

understanding of math concepts and procedures that correspond to overall mathematical achievement. To systematically support students, researchers have proposed implementing a multi-tiered system of supports (MTSS) framework, where school-based professionals provide high-quality, research-based academic, behavioral, and social-emotional interventions, systematically increasing supports from universal interventions for all students (Tier 1) to targeted (Tier 2) or individualized (Tier 3) interventions as a function of student need (Freeman et al., 2015; Eagle et al., 2015). Within this framework, response to intervention (RtI) involves the implementation of evidence-based instructional strategies linked with universal screening and ongoing progress monitoring to systematically identify and address gaps in students' academic skills (Bradley et al., 2005; Fuchs & Fuchs, 2006). Previous research supports the feasibility and effectiveness of several math interventions delivered within a RtI framework (Coddling et al., 2017). However, only 59% of elementary schools report implementing a RtI framework for mathematics (Spectrum K-12, 2010). These are relatively low levels of implementation, considering that 90% of elementary schools report RtI for reading (Spectrum K-12, 2010).

Of many barriers that may be contributing to these gaps in implementation, time may be the most significant (Mason et al., 2019). Intervention time blocks may be devoted to remediating reading problems, leaving math difficulties unattended. To maximize efficiency, it is important to consider the intervention duration and frequency per session or per week to identify the optimal strength (Coddling & Lane, 2015; DeFouw et al., 2018). For example, some students require an intervention two times a week, while other students require the same intervention four times a week to produce desirable outcomes (Coddling et al., 2017). The total amount of time the intervention occurs, in weeks or months (i.e., treatment duration) may be recommended for a specific period (e.g., 6–12 weeks; Gersten et al., 2009). However, if a student does not make adequate progress, the duration of the intervention may be lengthened (e.g., increase the number of total weeks). Finally, within a given intervention session, the number of minutes the intervention is delivered may be adjusted (Warren et al., 2007). Students struggling to become fluent in math computation may benefit from additional time to practice basic math facts. The session length of an intervention session can be intensified by increasing the amount of time (in min) a student is exposed to the intervention each session. Increasing the session length may allow for more opportunities (i.e., teaching episodes) for the student to practice the skill during an intervention session. The concept of adjusting time components (e.g., duration and frequency) of an intervention is known as treatment intensity.

Treatment intensity was adapted from the medical model, and within the field of education, some researchers define this concept as treatment strength, or the (a) dose, (b) dose form, (c) dose frequency, and (d) total treatment duration (Warren et al., 2007; Yoder & Woynaroski 2015). Other researchers perceive intensity as a broader construct conceptualized within a MTSS model to include aspects beyond treatment strength (i.e., treatment complexity, treatment expense; Barnett et al., 2004; Coddling & Lane, 2015; DeFouw et al., 2018; Mellard et al., 2010). This is distinct from, though related to, treatment fidelity, which is focused on how the intervention, once the treatment intensity is articulated (e.g., intervention to be delivered two times per week for 30-min), is implemented (Warren et al., 2007;

DeFouw et al., 2018). Although there are differences in how treatment intensity is conceptualized, most researchers agree that it is crucial to examine treatment intensity, or “in what form does the intervention work best” (Coddling & Lane, 2015, p. 2; Yeaton & Sechrest, 1981).

Treatment Intensity Within a RtI Framework

To implement RtI for mathematics, educators must determine how to efficiently intensify effective research-based services between tiers. However, there are clear inconsistencies between recommendations for how tier 2 and tier 3 mathematics interventions should be delivered in relation to dose form (i.e., group size; e.g., individual tutoring; Fuchs et al., 2008; small group; Bryant et al., 2008), session length (e.g., 20–30 min; Fuchs et al., 2008; 20–40 min; Gersten et al., 2009; Hunt, 2014), and dose frequency (e.g., three sessions per week; Fuchs et al., 2008; four to five sessions per week; Hunt, 2014). The Institute of Education Sciences (IES) RtI Mathematics Practice Guide (Gersten et al., 2009) recommends a specific dose form (i.e., small group), session length (i.e., 20–40 min), dose frequency (i.e., four to five times per week), and total treatment duration (i.e., 6–12 weeks). However, it is often not feasible for schools to implement tier 2 math intervention to these guidelines due to resource allocation, time constraints, and personnel availability concerns (Mason et al., 2019).

Researchers agree that the intervention, or treatment, should be intensified in terms of dose form, session length, dose frequency, and total duration (Coddling & Lane, 2015; Yoder & Woynaroski, 2015). However, specific recommendations for how interventions should be intensified between tiers are unclear, inconsistent, and lack empirical support. For research-based math interventions targeting students at-risk (e.g., below the 25th percentile), researchers most often report treatment strength aspects such as session length (number of minutes per session), dose frequency (number of sessions per week), and total treatment duration (number of weeks) with sessions averaging 25-min, three times per week for about 10 weeks (DeFouw et al., 2018). The number of teaching episodes (i.e., number of learning opportunities; e.g., math problems), which likely differ based on a multitude of variables (e.g., type of intervention, intervention time, student baseline levels), is often not reported in the math literature (DeFouw et al., 2018) making it difficult to understand *how many* learning opportunities are required during an intervention and how these learning opportunities are linked to student outcomes.

When school-based practitioners consult the literature on how to deliver research-based interventions within a RtI framework, uncertainty may arise when determining how to differentiate the intensity of math interventions between tier 2 and tier 3 (DeFouw et al., 2018). That is, for example, practitioners will not find clear guidance for how many minutes, how many students, or how often a research-based math intervention at tier 2 should be delivered. In practice, intervention characteristics (e.g., session length, size of group, frequency) will vary based on the students' needs (Duhon et al., 2009); however, evidenced-based guidelines provide a critical starting point to support the delivery of efficient and effective intervention service delivery.

To clarify this uncertainty, there is a need to further investigate and document characteristics of mathematics interventions to inform RtI research and practice.

Evaluations of Treatment Intensity

Researchers have begun to systematically evaluate aspects of treatment intensity in relation to intervention effectiveness (e.g., Coddling et al., 2017; Duhon et al., 2009). Specifically, researchers have examined manipulating treatment intensity for math interventions, in terms of varying dose frequency, to understand its impact on student outcomes.

Duhon and colleagues (2009) examined how to systematically increase dose frequency, or the number of sessions per day or week, in a math fact fluency class-wide intervention. Students were identified as high responders, average responders, and non-responders ($n=3$ students; Duhon et al., 2009). Using a multiple-baseline design, non-responder students received the intervention for a dose frequency of five times or 10 times a day (Duhon et al., 2009). Two of the three non-responders immediately increased their digits correct per min (DCPM) in the intensified five-times-per-day dose frequency intervention condition in comparison with previous class-wide intervention levels of DCPM (i.e., one-time-per-day dose frequency). The third non-responder increased DCPM in the intensified 10-times-per-day dose frequency condition. This study suggests that increasing dose frequency can be an effective intervention intensifier to support students who are not responding to class-wide interventions (Duhon et al., 2009).

Also evaluating treatment intensity, Coddling and colleagues (2016) examined the dose frequency (i.e., number of sessions per week) of a whole number math fluency small group packaged intervention while holding other aspects of treatment intensity constant (e.g., session length, duration). The session length was less than 20-min per session and the total treatment duration was 4 weeks, below the IES recommendation (i.e., 20–40 min; 6–12 weeks; Gersten et al., 2009). Students were randomly assigned to four dose frequency conditions: (a) control, (b) once a week, (c) twice a week, and (d) four times a week (Coddling et al., 2016). The packaged intervention involved three components: (a) guided practice, (b) independent practice, and (c) applied problems (Coddling et al., 2016). All dose frequency conditions (i.e., $1x$ /week, $2x$ /week, $4x$ /week) outperformed the control condition (Coddling et al., 2016). Further, students in the four-times-per-week dose frequency condition outperformed all groups on the math CBM measure, suggesting that the number of times the intervention occurred matters (Coddling et al., 2016).

Beyond the two dose frequency math intervention studies, there is also research comparing massed or distributed practice. Research found that distributed practice (i.e., four times per week; two times per week) significantly improved student math fluency growth rate compared to massed practice (i.e., once per week; Schutte et al., 2015). However, Coddling et al. (2019) found that opportunities to respond (OTR), or learning opportunities, and instructional level were a better predictor compared to massed or distributed spacing during a Cover-Copy-Compare intervention.

In sum, the current evidence suggests that there is a greater likelihood of producing desired student math fluency outcomes when interventions are more frequently delivered. None of these studies examined how dose frequency impacted student impressions or perception of the intervention. Further, we have not found empirical math intervention studies that have examined other aspects of treatment intensity such as session length.

Purpose of Study

Given the numerous demands placed on schools and resource constraints to implement math interventions, it is critical to understand the most time-efficient intervention session to best supports student outcomes. Session length, or how long the intervention should be delivered (in min) might offer a plausible mechanism for intensification. A systematic review of math interventions studies found that the average session frequency was below the IES recommended session length (DeFouw et al., 2018) and math interventions delivered for fewer than 20-min have demonstrated positive student math outcomes (e.g., Coddling et al., 2016). Thus, this study aimed to provide additional insight regarding how to conceptualize treatment intensity examining if an intervention delivered under the recommended guideline (i.e., 10-min) builds math fact fluency at the same rate compared to longer session lengths.

Specific research questions follow:

1. Does a whole number computation fluency intervention delivered at varying session lengths (0-min or control condition, 10-min, 20-min, 30-min) produce different DCPM for elementary students identified as at-risk for potential math failure when dose frequency and total treatment duration are held constant?
2. Do students rate the math intervention as assessed by the Kid's Intervention Profile (Eckert et al., 2017) as acceptable?

Method

Participants and Setting

Four first-grade students, identified as at-risk for mathematics failure, from a diverse, urban Title 1 charter elementary school in Massachusetts were selected to participate in this study. The school served approximately 650 students with majority of students identifying as African American (53.4%), followed by Hispanic (21.5%), and White (16.9%) (Massachusetts Department of Elementary and Secondary Education, 2019). Students at-risk for mathematics failure were identified using a Curriculum-Based Measurement (M-CBM; Shinn, 1989). Students who scored below 20 DCPM (Poncy & Duhon, 2015) on a grade-level M-CBM measure with less than 4 errors were eligible for this intervention. None of the students received special education services in mathematics. Parent consent and student assent were secured.

Student 1 was a 6-year-old Black female. Student 2 was a 7-year-old White female. Student 3 was a 6-year-old multiracial (i.e., Black and White) female. Student 4 was a 7-year-old Black female.

The primary researcher was a female, White doctoral school psychology student with training in research methods, school-based interventions, and data-based decisions. The primary researcher had previous experience with data collection and implemented all intervention sessions. The secondary rater, a female, Black doctoral school psychology student was trained to conduct direct observations of the intervention to ensure treatment fidelity and score CBM for interobserver agreement (IOA). Prior to the study, the primary researcher met with the secondary rater to review the purpose of the study, ask clarifying questions, discuss how to collect data and intervention procedures, and practice scoring CBM measures within a 1-h session. Both the primary researcher and secondary rater used answer keys to score the CBM measures; 100% IOA was reached during each session.

The study took place in a small group setting within a quiet empty classroom three times a week (one intervention session a day) and for a total treatment duration of 5 weeks. The session length was individualized for each student during each session. For example, while Student 1 received the 10-min intervention condition, Student 2 and Student 3 received the 30-min intervention condition, and Student 4 received the 20-min intervention condition during the same intervention block in a small group setting. During intervention, students were seated at separate desks and given all study materials (i.e., CCC worksheets, word problem worksheets, CBM assessment probe, and the KIP) in a folder. During the first session, the primary researcher taught the students to access target materials from the day from the left side of the folder, place them face down on their desk until told to start working, and flip them over when told to start. Students were taught to place completed materials in the right side of the folder. During intervention, students could not see the timer. When one or more students completed intervention earlier than their peers (based on the session length condition they were assigned to that day), they engaged in a preferred independent, non-academic task (e.g., putting stickers on their folder, drawing) at their desk until all peers completed their intervention time.

Materials

Materials included assessment math probes, intervention worksheets (i.e., CCC and WP), a stopwatch, acceptability measure, and integrity checklists. For assessment math probes, pre-assessments, baseline, intervention progress monitoring, and post-assessments were researcher-created M-CBMs. All assessment math probes included two pages of 35 math problems with five rows and seven columns (Poncy et al., 2007). Previous research yields adequate psychometric evidence for M-CBM tools (Burns et al., 2006; VanDerHeyden & Burns, 2008). Delayed alternate-form reliability for the probes is 0.85 when compared to similar tools (Foegen et al., 2007). Additionally, delayed alternate-form reliability estimates are higher for fluency single-skill probes (0.71) when compared to accuracy single-skill probes (0.49) (Burns et al., 2006).

Pre- and Post-Assessment

Prior to baseline and the intervention, two pre-assessment M-CBM probes were administered to assess multiple first-grade-level math skills: (1) *Addition to 20* fact families and (2) *Subtraction to 18* fact families. Following the pre-assessment probe, subskill mastery curriculum-based assessment (CBA) measures were created by the researcher to determine the students' appropriate instructional level as a survey-level assessment (SLA). This measure specifically measures a subdivision of broader grade-level skills (Hosp et al., 2007; Shapiro, 2010). For first grade, four single-skill CBA probes were created based on recommendations from VanDerHeyden and Burns (2008): (1) *Subtraction from 18*, (2) *Addition to 20*, (3) *Subtraction from 10*, and (4) *Addition to 10*. The purpose of the SLA was to identify specific math skills to target during the intervention. Following the intervention, two post-assessment M-CBM probes were administered to assess multiple first-grade-level skills: (1) *Addition to 20* fact families and (2) *Subtraction to 18* fact families.

Baseline Assessments and Progress Monitoring

Once a specific subskill was identified by the SLA, the researcher created four distinct math problem sets that corresponded to each condition (A, B, C, D) to monitor student progress throughout the intervention. The purpose of four distinct math problem sets was to ensure that math problems targeted during the four session length conditions were equivalent across sessions without overlap. All problem sets were created using procedures from Poncy et al. (2007, 2010) to ensure that each math problem set was mutually exclusive for each of the four session length conditions (10-min, 20-min, 30-min, control) and identical in presentation.

Basic subtraction facts were divided into four math problem sets containing seven problems each. Condition A (10-min) included: 8–4, 9–6, 10–8, 5–3, 9–3, 10–2, 5–2. Condition B (20-min) included: 6–3, 10–6, 7–5, 8–6, 10–4, 7–2, 8–2. Condition C (30-min) included: 4–2, 9–5, 10–7, 7–3, 9–4, 10–3, 7–4. Condition D (0-min) included: 10–5, 6–4, 8–5, 9–7, 6–2, 8–3, 9–2. All four problem sets contained the commutative fact families (e.g., “9–5” and “9–4”) and math problem answers in each problem set summed to 28. These procedures were consistent with previous recommendations (e.g., Poncy et al., 2007, 2010).

All assessments included two pages of 35 horizontal problems each. The worksheets included seven rows of five problems each. Each assessment probe (for progress monitoring) was created by first assigning a number to each math problem in the math problem set. For example, in Condition A (10-min), the first fact (i.e., 8–4) was assigned as “1” and the second fact (i.e., 9–6) was assigned as “2.” Next, assessment probes were developed using a random number generator to determine the presentation order of facts to ensure that each progress monitoring assessment probe was unique in presentation. Each progress monitoring assessment probe only included the math targets that corresponded to the session length condition. These procedures were consistent with previous recommendations (e.g., Poncy et al., 2007, 2010).

Kids Intervention Profile (KIP)

Social validity, or treatment acceptability, for the intervention was assessed using The Kids Intervention Profile (KIP; Eckert et al., 2017). The KIP is an 8-item scale that assesses students' perceptions of academic interventions (Eckert et al., 2017). The internal consistency of the KIP has been demonstrated to be 0.78 (Eckert et al., 2017). The KIP was developed using a 5-point anchored scale (*not at all to very, very much or never to many, many times*) with boxes to accompany Likert scale responses (Eckert et al., 2017). For this study, the questions were read to each student since the readability of this instrument is at a third-grade reading level (Eckert et al., 2017). The KIP was previously examined on a writing academic intervention (Eckert et al., 2017). The items were modified for a math to assess the acceptability of the intervention (see Appendix A).

Dependent Measures

The primary dependent variable was DCPM, which were computed by dividing the total number of correct digits by 2-min, for the assessment (pre, baseline, SLA, post) and progress monitoring. A digit was scored as correct when the correct number is written by the student in the proper column (Shinn, 1989). For example, given the problem “3 + 8” the student would receive two digits correct if “11” was provided. If the number “10” was written, the student would receive one digit correct for the “1” in the tens place. The second dependent variable was overall acceptability scores of the intervention on the KIP measure (Eckert et al., 2017).

Independent Variable

Session length was the independent variable and was operationally defined as four conditions: (a) *A* = 10-min, (b) *B* = 20-min, (c) *C* = 30-min, (d) control = 0-min. Two of these conditions (i.e., 20-min, 30-min) fell within the IES practice guideline recommending 20–40-min sessions. Across all conditions, the 10-min, 20-min, and 30-min length only accounted for the amount of time students were exposed to the math intervention. These times did not include transitions (e.g., students removing or placing materials in their folders) or progress monitoring. A timer was started as soon as the students began the intervention.

Across all students, the total session (i.e., intervention session, transitions, and progress monitoring) lasted a mean of 33:33 (SD = 13.12) minutes during baseline (progress monitoring for conditions *A*, *B*, *C*, *D* = 8:00), a mean of 21:32 (SD = 0.15) minutes during the 10-min condition (intervention session = 10:00, transitions = 9:32; progress monitoring = 2:00), a mean of 28:49 (SD = 0.28) minutes during the 20-min condition (intervention session = 20:00, transitions = 6:49, progress monitoring = 2:00), and a mean of 37:42 (SD = 0.07) minutes during the

Table 1 Average duration by session length condition

	Intervention	Progress monitoring	Transitions	Total time
Baseline (A, B, C, D)	–	8:00	–	33:33
10-min (A)	10:00	2:00	9:32	21:32
20-min (B)	20:00	2:00	6:49	28:49
30-min (C)	30:00	2:00	5:42	37:42
Control (D)	–	2:00	–	33:20

Numbers reflect minutes and seconds

30-min condition (intervention session = 30:00, transitions = 5:42, progress monitoring = 2:00). See Table 1 for more details.

Transition times were (a) time between intervention components (i.e., CCC and word problems), and (b) time following the intervention and start of progress monitoring. Baseline and intervention conditions (intervention session, transitions, and progress monitoring) aimed to occur for approximately 30-min. Transition times varied across conditions to account for differences in session length condition times. Progress monitoring times were consistent across conditions. The secondary rater recorded all times for the total session (i.e., intervention session, transitions, progress monitoring) for agreement.

Intervention

The first component of the intervention was a self-guided practice with basic math facts, or Cover-Copy-Compare (CCC), for half of the intervention session. The second intervention component was a guided practice with word problems for the other half of the intervention session. The IES practice guide (Gersten et al., 2009) recommends that Tier 2 math interventions should target basic math facts (i.e., CCC) supplemented with words problems to allow students to apply previously learned content. The same basic facts targeted during CCC were also targeted during the word problems intervention component. The total number of math problems (CCC + word problems) or teaching episodes were not held constant during the intervention across session length conditions. Across all students,

Table 2 Teaching episodes per minute by session length condition

	Student 1	Student 2	Student 3	Student 4	All students
10-min (A)	3.9	3.8	4.9	3.8	4.1
20-min (B)	3.4	2.9	4.8	3.3	3.6
30-min (C)	3.3	2.9	5.7	3.2	3.8

Teaching episodes = number of math problems (CCC + WP) completed during intervention block per min

the number of teaching episodes per minute (i.e., CCC + word problems) was 4.1 for 10-min condition, 3.6 for the 20-min condition, and 3.8 for the 30-min condition. See Table 2 for more details.

Cover-Copy-Compare

Students participated in self-guided practice for basic math facts using a traditional CCC procedure. Research suggests that CCC is an appropriate intervention selection for students who have an accuracy above 60% and below 20 DCPM on the M-CBM screening tool (Poncy & Duhon, 2015). There were five steps for CCC: (1) student looked at math problem and answer; (2) covered the math problem; (3) copied and answered the math problem beside the covered one; (4) uncovered the first math problem; and (5) compared response to initial answer (Coddling et al., 2009). If the student writes the correct problem, they move to the next math problem. If the student writes the incorrect problem, they repeat the five CCC steps. Each CCC worksheet contained one column of single-skill problems with the answer provided on the left side of the paper. There was a dotted line to indicate where to fold the paper to cover the answer. Then, the student copied the math problem with the answer on the right side of the paper. Students completed the five steps described for each of the problems. During the first session, the primary investigator explicitly explained the steps of CCC to the students as a small group, modeled the intervention, and provided an opportunity for students to practice the strategy. Each student was able to demonstrate all the steps of the intervention independently following instruction.

Schema-Based Word Problem Instruction

The students participated in schema-based applied mathematics word problems instruction (Coddling et al., 2016; Jitendra, 2016; VanDerHeyden et al., 2012). Researcher-created applied problem worksheets included grade-level word problems (Coddling et al., 2016). During this intervention component, word problems were initially taught during the first intervention session using a schema-based instruction by the primary researcher with students (Jitendra, 2007), aligning with WWC guidelines for math interventions (Fuchs et al., 2021). After students were taught schema-based instruction, they independently completed the worksheets as self-guided practice. Schema-based instruction teaches the student to organize information into a framework of known and unknown information (Jitendra, 2007). The purpose of this instruction allows the student to identify, represent, and solve the word problem. The researcher modeled and taught student to complete the following 5 steps using schema-based instruction: (1) read word problem, (2) identify relevant from irrelevant information, (3) represent the problem correctly (change, group, compare), (4) select appropriate strategy for solving the problem, and (5) solve the problem computational. Students were provided with three problem types that were randomly administered across each condition: (a) change (e.g., Susan has 8 flowers, and then she picks 2 more flowers. How many flowers does she have now?); (b) group (e.g., Juan is playing basketball with his friends; if he has 2 friends on his team and 3 friends on the other team, how many friends is Juan playing basketball with?); or (c)

compare word problem (e.g., there are 25 students in the class and only 10 have a pencil. How many students need a pencil?) (Jitendra, 2007). Math problems targeted during the CCC intervention component were the same math problems embedded into word problems.

Intervention Conditions

Condition A occurred for 10-min with 5-min allocated for CCC and 5-min for word problems. Following the intervention session, students completed the 2-min M-CBM assessment probe targeting the math problem Set A following the 10-min intervention.

Condition B occurred for 20-min procedure with 10-min allocated for CCC and 10-min for word problems. Students completed the 2-min M-CBM assessment probe targeting the math problem Set B following the 20-min intervention.

Condition C occurred for 30-min with 15-min allocated for CCC and 15-min for word problems. Following the intervention session, students completed the 2-min M-CBM assessment probe targeting the math problem Set C following the 30-min intervention.

Control condition included students completing a 2-min M-CBM assessment probe for the math problem Set D. Student did not receive the intervention. Following the M-CBM assessment probe, students engaged in non-related math activities until the other students in the small group completed their intervention sessions (e.g., reading a book, playing a game).

Procedures

After obtaining university IRB approval, the researcher sought principal approval and interest to support struggling students with basic math computation fluency. The study included screening, baseline, and intervention phases to compare varying session lengths on student mathematics performance. Student 2 received a verification phase, condition that produced the greatest DCPM during the intervention phase. Due to the end of the school year, a verification phase was not implemented for Students 1, 3, and 4.

Screening and Survey-Level Assessment

Prior to the intervention, the researchers administered the screening class-wide identifying four students as eligible for this study. Following the screening, the pre-assessment M-CBM assessment was administered. Next, survey-level assessments (SLA) were individually administered to students outside of the classroom to determine each student's appropriate instructional level and specific mathematics computation skills to target during the intervention (Hosp et al., 2007; Shapiro, 2010). The SLA was conducted in a small group with the four eligible participants based on the initial M-CBM assessment probe class-wide. The skill that had an accuracy above 60% and below 20 DCPM served as the target skill for the intervention meeting

CCC requirements (Poncy & Duhon, 2015). Following the SLA, all students scored above 20 DCPM for *Addition to 10* and below 20 DCPM for *Subtraction from 10* (Student 1 DCPM=7.5, accuracy=100%; Student 2=DCPM=9, accuracy=100%; Student 3=DCPM=6, accuracy=90%; Student 4 DCPM=4, Accuracy=80%).

Baseline

During baseline, the researcher met with participants in a separate room to complete a 2-min assessment M-CBM probe on three separate occasions on all four math problem sets (A, B, C, D; Kratochwill et al., 2010). No intervention was implemented.

Alternating Treatment Phase

After baseline data were gathered, the order in which each student would receive the four session length conditions ($A=10$ -min, $B=20$ -min, $C=30$ -min, control= 0 -min) was randomly assigned using a random number generator (e.g., Student 1's sequence for five repetitions: ACB; BCA; CBA; CAB; ABC). The intervention was delivered in a small group setting; however, the intervention sequence was individualized for all students. For example, during the first day of the treatment phase for Student 1, the intervention package was delivered for 10-min (5-min for CCC and 5-min for word problems) using the problems that were assigned to Math Problem Set A, followed by a progress monitoring assessment probe targeting math problems in Set A, non-related math activities for the remaining intervention block (approximately 20-min). However, for Student 2, during the first day of the treatment phase, the intervention package was delivered for 30-min (15-min for CCC and 15-min for word problems) using the problems that were assigned to Math Problem Set C, followed by a progress monitoring assessment probe targeting math problems in Set C. Following the intervention and progress monitoring assessment probe, student completed the KIP (social validity rating) for the current intervention block. Students were told to complete the KIP based on the day's session and not to rate based on other intervention sessions. The researcher delivered the packaged intervention three times a week for 5 weeks.

Procedural Fidelity

A checklist was created to ensure that the researcher delivered each component of the intervention. Adherence to each step of the intervention procedures was assessed as fully implemented or not implemented. The primary researcher indicated the duration of each intervention component to verify that each condition was delivered accordingly to the allotted amount of time. To ensure that the intervention was implemented as intended, the secondary rater also completed a treatment fidelity checklist. The secondary rater observed an average of 31.4% of sessions (baseline=33, 10-min=23.5%; 20-min=35.3%; 30-min=58.8%; control=41.7%) for 100% adherence across all conditions.

Interobserver Agreement

The secondary rater scored at least 35% of all intervention materials including the M-CBM assessment probes, CCC worksheets, and word problem worksheets. Agreement was calculated by dividing the number of agreements by the total number of agreements and disagreements. On average, the secondary rater scored 38.5% of M-CBM assessment probes (baseline = 33%, *A* = 23.5%, *B* = 35.3%, *C* = 58.8%, control = 41.7%) for an average of 96.7% agreement (baseline = 100%, *A* = 100%, *B* = 83.3%, *C* = 100%, control = 100%). The secondary rater observed and scored on average 41.5% of CCC worksheets across all conditions (Student 1 = 40.0%; Student 2 = 42.9%; Student 3 = 41.7%; Student 4 = 41.7%) with 100% agreement across all participants and conditions indicating high IOA. The secondary rater observed and scored on average 41.5% of word problem worksheets across all conditions (Student 1 = 40.0%; Student 2 = 42.9%; Student 3 = 41.7%; Student 4 = 41.7%) for 90.9% agreement across all participants and conditions indicating acceptable IOA.

Experimental Design and Data Analysis

The current study used an adapted alternating treatment design to allow two or more conditions to be simultaneously compared through equivalent and independent instructional items (i.e., problem sets) with a control condition accounting for history and spillover effects (Cooper et al., 2019; Kratochwill et al., 2014). To decrease student's frequency of inaccurate responding and frustration, the control condition was assessed following the three intervention conditions (Christ, 2007; Cuvo, 1979). The order of the interventions (i.e., 10-min, 20-min, 30-min sessions) was randomized with at least four repetitions of the alternating sequence (Kratochwill et al., 2010) across sessions while duration was held constant, that is, the number of intervention weeks were the same for all conditions and students (e.g., 5 weeks). The control condition (0-min) occurred following the end of an alternating sequence (e.g., ABC D).

For adapted alternating treatment designs, examining changes in trend, level, and variability of data paths is recommended and considered suitable for this specific design (Kratochwill et al., 2010; Manolov & Onghena, 2018). As such, data paths within session length conditions and across time for specific outcome measures were visually analyzed (Kratochwill et al., 2010). To supplement visual analysis, means, standard deviations were calculated for each phase and condition. Rate of improvement ($ROI = \frac{\text{Post-Assessment Score} - \text{Pre-Assessment Score}}{\# \text{ of intervention weeks}}$), which is compared to expected growth rate (Solomon et al., 2020), was calculated for each student. Social validity data were summarized (i.e., means and standard deviations) to determine students' perceptions of the overall intervention. Additionally, this study met What Works Clearinghouse (WWC) evidence standards with reservations for an alternating treatment design by: (a) systematically manipulating the independent variable; (b) assessing interobserver agreement across at least 20% of measures and sessions for all phases;

(c) having at least four points per condition with at least two points per phase. However, this study met WWC evidence standards with reservations because it did not have five points per condition (WWC, 2020) due to end of the school year (Fig. 1).

Table 3 Mean (SD) digits correct per minute (DCPM) across session length condition

Participant	Baseline				Intervention			
	10-min (A)	20-min (B)	30-min (C)	Control (D)	10-min (A)	20-min (B)	30-min (C)	Control (D)
Student 1	6.2 (1.04)	4.5 (0.87)	6.0 (1.80)	7.0 (3.50)	15.4 (2.43)	16.3 (3.17)	13.8 (3.70)	8.1 (3.93)
Student 2	8.3 (0.29)	8.0 (0.87)	8.0 (2.00)	8.2 (2.78)	13.4 (4.37)	13.8 (1.89)	16.6 (3.47)	9.9 (2.78)
Student 3	6.3 (0.58)	7.2 (4.48)	7.3 (0.76)	9.5 (0.87)	19.4 (2.50)	19.3 (6.96)	20.0 (3.69)	14.9 (3.04)
Student 4	6.5 (1.80)	5.8 (2.02)	8.0 (2.29)	4.3 (0.76)	17.6 (2.10)	14.1 (4.33)	14.9 (2.50)	9.25 (1.85)

Table 4 Total mean score social validity by session length condition across students

Participant	10-min (A)	20-min (B)	30-min (C)	Total intervention
Student 1	40.00	40.00	40.00	40.00
Student 2	39.00	39.50	39.50	39.43
Student 3	37.75	37.50	35.25	36.83
Student 4	40.00	40.00	40.00	40.00
All students	35.83	36.17	36.17	36.26

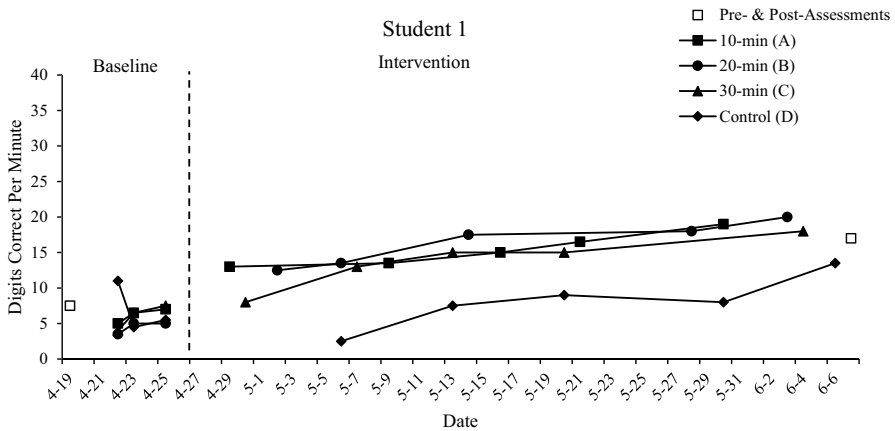


Fig. 1 DCPM by session across conditions

Results

Findings are described as they relate to research questions. First, session length and DCPM data are presented (see Table 3). Second, social validity data are presented (see Table 4).

Session Length and DCPM

Student 1

During baseline, Student 1’s data were relatively stable at a low level for the 10-min condition ($M=6.2$ DCPM; $SD=1.04$ DCPM), the 20-min condition ($M=4.5$ DCPM; $SD=0.87$ DCPM), the 30-min condition ($M=6.0$ DCPM; $SD=1.80$ DCPM), and the control condition ($M=7.0$ DCPM; $SD=3.50$ DCPM; see Fig. 2). All three conditions displayed a slight increasing trend, and all three conditions were differentiated with the control condition ($M=8.1$ DCPM; $SD=3.93$ DCPM), favoring the intervention conditions. No differentiation was observed among any of the conditions, illustrating that the student performed similarly across all four math problem sets during baseline.

During the intervention phase, student 1 performed similarly among all three treatment conditions with a mean of 15.4 DCPM ($SD=2.43$ DCPM) for the 10-min condition, mean of 16.3 DCPM ($SD=3.17$) for the 20-min condition, and a mean of 13.8 DCPM ($SD=3.70$) for the 30-min condition. A level change was observed with the introduction of the intervention phase for the 10-min and 20-min condition but not for the 30-min condition. Additionally, prior to the intervention, Student 1’s pre-assessment was 7.5 DCPM. Following the intervention, Student 1’s post-assessment increased to 17 DCPM ($\Delta=10.5$). Overall, Student 1’s ROI on the M-CBM measure from pre-assessment to post-assessment was 2.1 DCPM, which is greater than the average expected growth rate of 1.62 (Solomon et al., 2020).

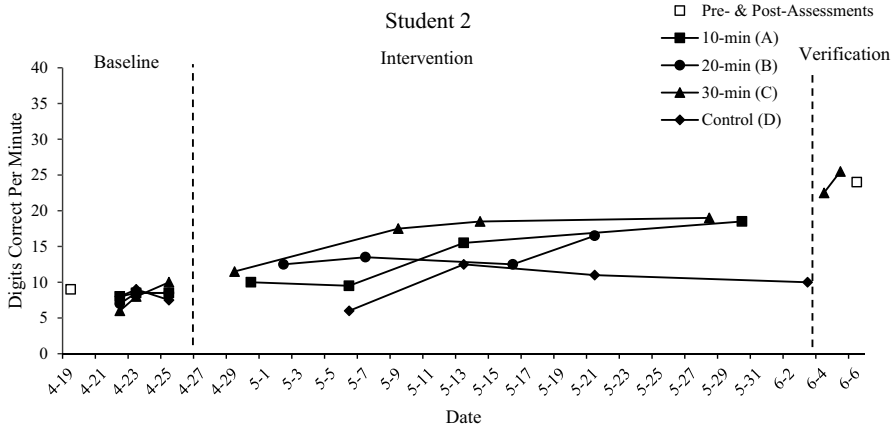


Fig. 2 DCPM by session across conditions

Student 2

During baseline, Student 2's performance was relatively stable at a low level for the 10-min condition ($M=8.3$ DCPM; $SD=0.29$ DCPM), the 20-min condition ($M=8.0$ DCPM; $SD=0.87$ DCPM), the 30-min condition ($M=8.0$ DCPM; $SD=2.00$ DCPM), and the control condition ($M=8.2$ DCPM; $SD=2.78$ DCPM; see Fig. 2). No differentiation was observed among any of the conditions, illustrating that the student performed similarly across all four math problem sets during baseline.

With introduction of the intervention, differentiation was observed between all three intervention conditions and the control ($M=9.9$ DCPM; $SD=2.78$ DCPM), favoring the intervention conditions. Performance under the 30-min ($M=16.6$ DCPM; $SD=3.47$ DCPM) was differentiated from the 10-min ($M=13.4$; $SD=4.37$) and 20-min condition ($M=13.8$ DCPM; $SD=1.89$ DCPM), favoring the 30-min condition. For the 30-min condition, a slight level change was observed as was an increasing trend. Performance in the final phase was higher than both baseline and intervention phases and increased over the two sessions. With the onset of the intervention phase, no level change was observed for the 10-min condition ($M=13.4$ DCPM; $SD=4.37$ DCPM) but an increasing trend was observed. No differentiation was observed with the 20-min condition. For the 20-min condition, a level change was observed with a stable trend that increased during the last treatment session ($M=13.8$ DCPM; $SD=1.89$ DCPM). Additionally, prior to the intervention, Student 2's pre-assessment was 9 DCPM. Following the intervention, Student 2's post-assessment increased to 24 DCPM ($\Delta=15$). Overall, Student 2's ROI on the M-CBM measure from pre-assessment to post-assessment was 3.0 DCPM, which is greater than the average expected growth rate of 1.62 (Solomon et al., 2020).

Student 3

During baseline, Student 3's data were relatively stable at a low level for the 10-min condition ($M=6.3$ DCPM; $SD=0.58$ DCPM), 20-min condition ($M=7.2$ DCPM; $SD=4.48$ DCPM), 30-min condition ($M=7.3$ DCPM; $SD=0.76$ DCPM), and the control condition ($M=9.5$ DCPM; $SD=0.87$ DCPM; see Fig. 3). There was slight differentiation of performance between the control and 20-min math problem sets with the 10- and 30-min math problem sets, but these differences are slight and not clinically meaningful, suggesting that Student 3 performs similarly across math problem sets.

With the onset of the intervention phase, no differentiation was observed among intervention conditions; however, performance under all three intervention conditions was higher than the control condition ($M=14.9$ DCPM; $SD=2.50$ DCPM). On average, during the 10-min condition, Student 3's DCPM were 19.4 ($SD=2.50$). An immediate level change was observed with the onset of the intervention phase and an increasing trend was observed. For the 20-min condition, Student 3's DCPM were on average 19.3 ($SD=6.96$) and an increasing trend was observed with no level change. For the 30-min condition, Student 3 produced an average of 20.0 ($SD=3.69$) DCPM with a level change and increasing trend. Additionally, prior to

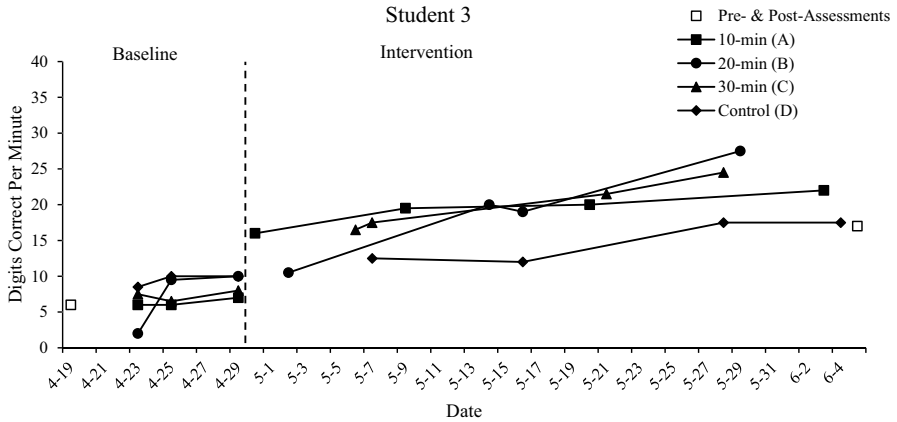


Fig. 3 DCPM by session across conditions

the intervention, Student 3’s pre-assessment was 6 DCPM. Following the intervention, Student 3’s post-assessment increased to 17 DCPM ($\Delta = 11$). Overall, Student 3’s ROI on the M-CBM measure from pre-assessment to post-assessment was 2.2 DCPM, which is greater than the average expected growth rate of 1.62 (Solomon et al., 2020).

Student 4

During baseline, Student 4’s data were relatively stable at a low level for the 10-min condition ($M = 6.5$ DCPM; $SD = 1.80$ DCPM), 20-min condition ($M = 5.8$ DCPM; $SD = 2.02$ DCPM), 30-min condition ($M = 8.0$ DCPM; $SD = 2.29$ DCPM), and the control condition ($M = 4.3$ DCPM; $SD = 0.76$ DCPM; see Fig. 4). No meaningful

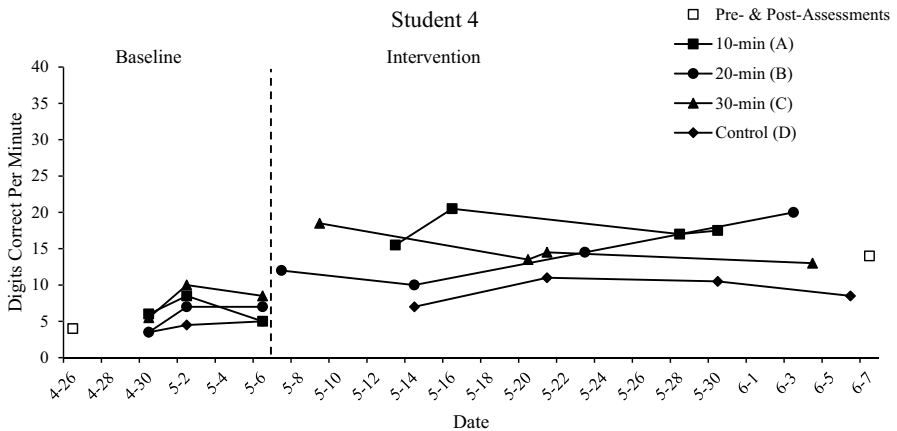


Fig. 4 DCPM by Session Across Conditions

differentiation was observed between the math problem sets during baseline, confirming that Student 4 performed similarly across math problem sets.

With the onset of the intervention phase, although slight separation was observed among all three intervention conditions, the patterns did not differ in relation to overall level, trend, and variability across the three conditions. Performance among all three conditions was higher than the control condition and all three conditions differentiated from the control condition ($M=9.25$ DCPM; $SD=1.85$ DCPM). With the onset of the intervention phase with the 10-min condition, an immediate level change was observed which initially increased and then returned to the level from the first session ($M=17.6$ DCPM; $SD=2.10$ DCPM). For the 20-min condition, an immediate level change and increasing trend was observed ($M=14.1$ DCPM; $SD=4.33$ DCPM). With the 30-min condition, Student 4 displayed an immediate level change followed by a decreasing trend ($M=14.9$ DCPM; $SD=2.50$ DCPM). Additionally, prior to the intervention, Student 4's pre-assessment was 4 DCPM. Following the intervention, Student 4's post-assessment increased to 14 DCPM ($\Delta=10.0$). Overall, Student 4's ROI on the M-CBM measure from pre-assessment to post-assessment was 2.0 DCPM, which is greater than the average expected growth rate of 1.62 (Solomon et al., 2020).

Social Validity

Students completed the KIP (Eckert et al., 2017) to indicate their perception of both session length conditions and overall intervention (See Table 4 for more details). Overall, Student 1 indicated that all intervention sessions were highly acceptable with a total mean score of 40.00. Student 1 also found the three session length conditions highly acceptable (10-min=40.00; 20-min=40.00; 30-min=40.00). Student 2 indicated that all intervention sessions were highly acceptable with a total mean score of 39.43. Student 2 also found the three session length conditions acceptable (10-min=39.00; 20-min=39.50; 30-min=39.50). Overall, Student 3 indicated that all intervention sessions were acceptable with a total mean score of 36.83. Student 3 found both the 10-min (37.75) and 20-min (37.50) conditions highly acceptable, and the 30-min condition acceptable (35.25). Student 4 indicated that all intervention sessions were highly acceptable with a total mean score of 40.00. Student 4 also found the three session length conditions highly acceptable (10-min=40.00; 20-min=40.00; 30-min=40.00).

Discussion

Despite the availability of numerous effective math interventions (Coddling et al., 2017), the percentage of schools implementing RtI to support math development and achievement continues to lag behind reading (Spectrum K-12, 2010). Implementation barriers including lack of change agents, access to resources, and time impede schools' ability to deliver math services feasibly and effectively within MTSS (Mason et al., 2019). Due to the lack of attention given to understanding

various aspects of treatment intensity for math specifically, there are inconsistencies between research and practice guidelines for how to implement these interventions (DeFouw et al., 2018). Currently, there are limited math interventions studies (e.g., Coddling et al., 2016; Duhon et al., 2009) that examined one aspect of treatment intensity—dose frequency. The purpose of this study was to contribute to the emerging treatment intensity literature, by focusing on how to efficiently deliver math interventions in terms of session length. The current recommendations for delivering interventions offer a large window of time for the length of single intervention sessions (i.e., 20–40 min; Gersten et al., 2009). This study systematically examined and manipulated the session length (i.e., the number of min per session) of a packaged math intervention. Overall, results suggest three major findings: (1) all three intervention conditions outperformed the control condition (0-min), (2) student outcomes in the 10-min condition were similar to outcomes in the longer 20- and 30-min conditions, and (3) intervention session length may be investigated case by case for individual students using a multi-tiered framework.

Impact of Session Length on Student Outcomes

In this study, all three treatment conditions outperformed the no-treatment control condition. Following intervention, all four students' performance under each intervention conditions exceeded the screening criterion (i.e., 20 DCPM) within the instructional range (14–31 DCPM; Burns et al., 2006). For three of the four participants (Students 1, 2, and 4), performance continued to fall in the frustrational range on the control math problem set. For Student 3, performance in the control condition fell in the instructional range at the end of the intervention phase (albeit still lower than the performance under the intervention conditions). This student had the highest rate of performance across all conditions compared to the other participants. Emerging research on expected growth rates in math suggests that the average growth across intervention sessions is 1.62 DCPM (Solomon et al., 2020). In the current study, all four students' growth rate was at or above the expected 1.62 DCPM.

Regarding session length, no distinct condition outperformed the other conditions for three of the four students (Students 1, 3, and 4). That is, most students in the 10-min condition performed similarly to the 20- and 30-min conditions. The IES practice guideline (Gersten et al., 2009) recommends 10-min of fact fluency practice across all tiers. The current study included fact fluency practice for fewer minutes than recommended (Condition A, 5-min), at the current recommendation (Condition B, 10-min), and more minutes than recommended (Condition C, 15-min). All students completed a similar rate of teaching episodes per minute across all session length conditions. Interestingly, for all students, the 10-min condition had the highest rate of teaching episodes completed, suggesting in the briefest condition, students were exposed to more problems per min. Given that we used a gated screening approach to isolate instructional planning to the specific skill area of need (identified by the SLA) and reduced the set size (Solomon et al., 2020), students were more successful with smaller units of time for the intervention (i.e., 10-min condition).

This study is one of the first in the math intervention literature to examine session length in isolation. Previous research examining dose frequency (i.e., number of sessions per week), found delivering a packaged math intervention four times weekly had the greatest benefit for students on a math computation proximal measure (Coddling et al., 2016). However, the four times weekly did not outperform less frequent conditions (i.e., once weekly, twice weekly) on other measures (e.g., application; Coddling et al., 2016). Frequency per session is one way to alter the time allocated to an intervention. The number of minutes of each intervention session can be adjusted depending on student response to intervention similarly to the number of intervention sessions per week. This preliminary data suggest that brief intervention sessions for young children on targeted math skills, like basic math facts, may be useful. This is an important finding because less time-intensive interventions are more feasible within schools where time is scarce. In this study, one student benefited from longer intervention sessions; however, it is unclear if more sessions per week would have also been useful to improve this student's performance. The study purpose did not permit evaluation between different aspects of treatment dose (i.e., Coddling et al., 2016; Duhon et al., 2009).

Finally, this study found individual differences between students in outcomes, indicating that intervention session length may be investigated case-by-case using a multi-tiered framework. For example, Student 2 performed best during the 30-min condition. As a result, this condition was continued during the verification phase, and this student's performance continued to improve until the intervention was terminated. Notably, Student 2 completed the highest rate of teaching episodes per min during the 10-min condition with similar rates in both the 20-min and 30-min conditions, suggesting that more teaching episodes may not be the factor influencing this student's performance. Overall, these findings indicate that for some students a shorter session length (i.e., 10-min) may result in similar subtraction fact fluency when compared to longer session lengths (i.e., 20-min and 30-min). The performance of all students fell within the frustration range of performance for subtraction during pre-testing, demonstrating 10 or fewer DCPM (Burns et al., 2006).

Students' perceptions of the math intervention were assessed following each intervention session using the KIP (Eckert et al., 2017). All students found the intervention acceptable and variability in ratings was minimal across intervention sessions. This finding is congruent with previous research that found that most interventions in school-based research are rated as acceptable (Silva et al., 2020). Specifically, students perceived that their math improved, did not worsen, and there were limited times when they did not want to practice math or participate in the intervention. Previous studies examining treatment intensity components (e.g., Coddling et al., 2016; Duhon et al., 2009, 2015) did not assess students' perception of the acceptability of the intervention. Therefore, little is known on how students perceive aspects of treatment intensity.

Limitations

There are several design and methodological limitations to consider when interpreting the findings of the current study. First, the total treatment duration was 5 weeks, and the dose frequency was three times per week—less than research (Coddling et al., 2016)

and IES tier 2 recommendations (Gersten et al., 2009), due to practical constraints and the school year ending. Student performance may have reached mastery if the intervention was delivered according to duration (i.e., 6–8 weeks) and frequency (i.e., four times per week) recommendations.

Second, the current study targeted math fact fluency using one intervention (i.e., CCC) and findings may not extend to other math interventions (e.g., Explicit Timing, Detect-Practice-Repair). Since fluency was the primary dependent variable, word problem-solving was not assessed. It is unknown how session length influenced word problem outcomes. Future research may explore how opportunities to practice (i.e., longer sessions) relate to generalization skills (e.g., word problem-solving). This study did not record the duration of initial teaching sessions for word problems, which could be documented in future research.

Third, this study did not collect student behavior data (e.g., engagement), which could provide more context regarding session length results. Fourth, all participants were female and selected from only one school. Future studies with a more diverse sample are warranted. Fifth, acceptability was assessed for the whole packaged intervention and looking at preferences regarding intervention components may be explored in future research. Teacher acceptability was not assessed but may be informative to future research.

Sixth, the condition control was assessed only three times. Future research may consider exploring equal intervention and control sessions. Additionally, although procedures were used to develop mutually exclusive probe sets, there was a potential for carryover effect for student learning across session length conditions. Finally, the number of learning opportunities (CCC and word problems) was not held constant across sessions. Future research may continue to document and explore the impact of teaching episodes when varying intervention components.

Implications

There is limited research studying how to understand and conceptualize treatment intensity in the math intervention (Coddling & Lane, 2015; DeFouw et al., 2018). This study contributes to understanding the treatment intensity aspect of session length. This study demonstrated that continual fluency growth was similar across all session length conditions for majority of the students (i.e., three out of four). During this 5-week targeted, supplemental small group intervention delivered three times a week for a range of session lengths from 10 to 30 min, all four students reached an instructional level (Burns et al., 2006). This finding suggests that an intervention session length less than the 20–40-min recommendation can improve students' math fluency to an instructional level. A shorter intervention session length (i.e., 10-min) may be similarly effective while being a more time-efficient way to intervene for students who are at-risk for math failure in schools with limited resources.

Conclusion

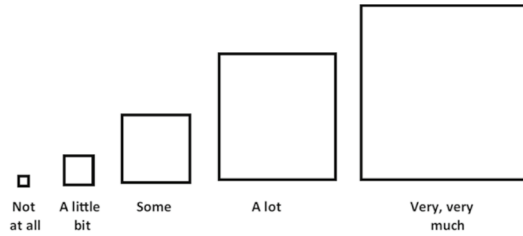
This study examined the impact of varying session lengths on a packaged math fact fluency intervention. Overall, results indicate that student fluency increased across the three session length conditions (10-, 20-, 30-min) compared to the control condition (0-min). These results suggest that shorter session lengths (i.e., 10-min) may be equally effective in improving students' math fluency as longer session lengths (i.e., 30-min). It is important to note that there were individual differences demonstrated between students on intervention session length suggesting that session length may be a case-by-case basis recommendation aligning with a multi-tiered framework. Finally, all students indicated that the intervention was acceptable. The results of the current study contribute to the growing literature for how to deliver a small group math intervention effectively and efficiently. Additional research is encouraged for further contribution to understanding how various treatment intensity aspects can support students' needs.

Appendix A

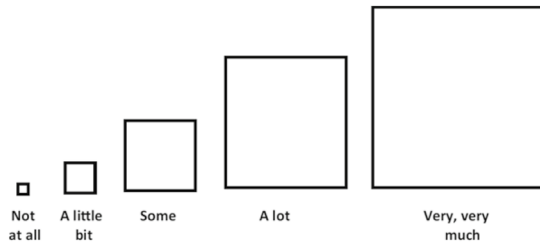
(Adapted from the Kids Intervention Profile (KIP); Eckert et al., 2017).

Directions: Please indicate how much you agree with the questions.

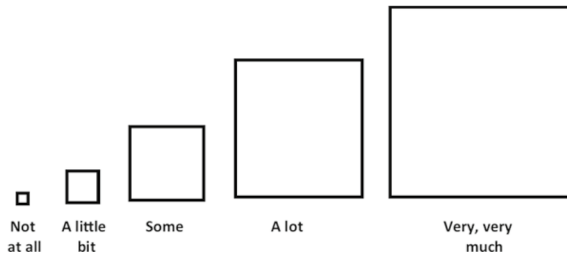
1. How much do you like practicing math with us each time?



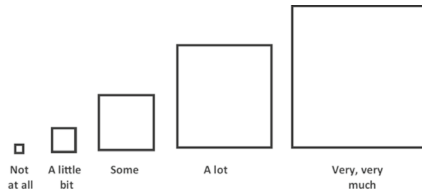
2. How much do you like being told what to do math about?



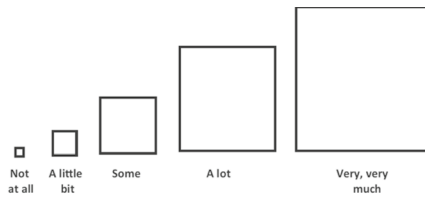
3. Were there times when you didn't want to practicing math with us?



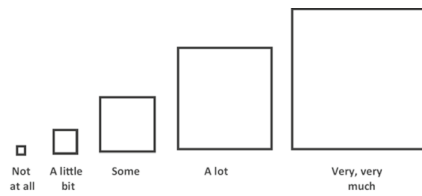
4. Were there any times when you wished you could work more on practicing math with us?



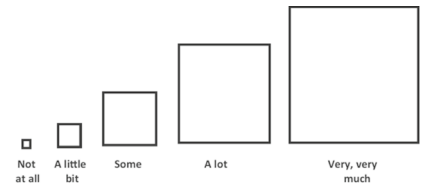
5. How much do you like being told how many math problems you to do?



6. How much do you think it helps you when you were told how many math problems you did?



7. Do you think your math has improved?



8. Do you think your math has gotten worse?

Declarations

Conflict of interest The authors have no conflicts of interests to declare.

Informed Consent Teacher and parent consent and student assent was obtained prior to the intervention.

References

- Barnett, D. W., Daly, E. J., Jones, K. M., & Lentz, F. E., Jr. (2004). Response to intervention: Empirically based special service decisions from single-case designs of increasing and decreasing intensity. *The Journal of Special Education, 38*(2), 66–79. <https://doi.org/10.1177/00224669040380020101>
- Bradley, R., Danielson, L., & Doolittle, J. (2005). Response to Intervention. *Journal of Learning Disabilities, 38*, 485–486. <https://doi.org/10.1177/00222194050380060201>
- Bryant, D. P., Bryant, B. R., Gersten, R. M., Scammacca, N. N., Funk, C., Winter, A., Shih, M., & Pool, C. (2008). The effects of tier 2 interventions on the mathematics performance of first-grade students who are at risk for mathematics difficulties. *Learning Disability Quarterly, 31*, 47–63. <https://doi.org/10.2307/20528817>
- Burns, M. K., VanDerHeyden, A. M., & Jiban, C. (2006). Assessing the instructional level for mathematics: A comparison of methods. *School Psychology Review, 35*, 401–418.
- Christ, T. J. (2007). Experimental control and threats to internal validity of concurrent and nonconcurrent multiple baseline designs. *Psychology in the Schools, 44*(5), 451–459. <https://doi.org/10.1002/pits.20237>
- Codding, R. S., Chan-Iannetta, L., Palmer, M., & Lukito, G. (2009). Examining a classwide application of Cover-Copy-Compare with and without goal setting to enhance mathematics fluency. *School Psychology Quarterly, 24*, 173–185. <https://doi.org/10.1037/a0017192>
- Codding, R. S., & Lane, K. L. (2015). A spotlight on treatment intensity: An important and often overlooked component of intervention inquiry. *Journal of Behavioral Education, 24*(1), 1–10. <https://doi.org/10.1007/s10864-014-9210-z>
- Codding, R. S., Volpe, R. J., Martin, R. J., & Krebs, G. (2019). Enhancing mathematics fluency: Comparing the spacing of practice sessions with the number of opportunities to respond. *School Psychology Review, 48*, 88–97. <https://doi.org/10.17105/SPR-2018-0010.V48-1>
- Codding, R. S., Volpe, R. J., & Poncy, B. C. (2017). *Effective math interventions: A guide to improving whole-number knowledge*. Guilford Press.
- Codding, R. S., VanDerHeyden, A. M., Martin, R. J., Desai, S., Allard, N., & Perrault, L. (2016). Manipulating treatment dose: Evaluating the frequency of a small group intervention targeting whole number operation. *Learning Disabilities Research & Practice, 31*, 208–220. <https://doi.org/10.1111/ldrp.12120>
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2019). *Applied Behavior Analysis* (3rd ed.). Pearson Education, Inc.
- Cuvo, A. J. (1979). Multiple-baseline design in instructional research: Pitfalls of measurement and procedural advantages. *American Journal of Mental Deficiency, 84*, 219–228.
- DeFouw, E. R., Codding, R. S., Collier-Meek, M. A., & Gould, K. M. (2018). Examining dimensions of treatment intensity and treatment fidelity in mathematics intervention research for students at-risk. *Remedial and Special Education, 40*(5), 298–312. <https://doi.org/10.1177/0741932518774801>
- Duhon, G. J., House, S., Hastings, K., Poncy, B., & Solomon, B. (2015). Adding immediate feedback to explicit timing: An option for enhancing treatment intensity to improve mathematics fluency. *Journal of Behavioral Education, 24*, 74–87. <https://doi.org/10.1007/s10864-014-9203-7>
- Duhon, G. J., Mesmer, E. M., Atkins, M. E., Greguson, L. A., & Olinger, E. S. (2009). Quantifying intervention intensity: A systematic approach to evaluating student response to increasing intervention frequency. *Journal of Behavioral Education, 18*, 101–118. <https://doi.org/10.1007/s10864-009-9086-5>
- Eagle, J. W., Dowd-Eagle, S. E., Snyder, A., & Holtzman, E. G. (2015). Implementing a multi-tiered system of support (MTSS): Collaboration between school psychologists and administrators to promote systems-level change. *Journal of Educational and Psychological Consultation, 25*(2–3), 160–177. <https://doi.org/10.1080/10474412.2014.929960>
- Eckert, T. L., Hier, B. O., Hamsho, N. F., & Malandrino, R. D. (2017). Assessing children's perceptions of academic interventions: The Kids Intervention Profile. *School Psychology Quarterly, 32*, 268–281. <https://doi.org/10.1037/spq0000200>
- Foegen, A., Jiban, C., & Deno, S. (2007). Progress monitoring measures in mathematics: A review of the literature. *The Journal of Special Education, 41*, 121–139. <https://doi.org/10.1177/00224669070410020101>

- Freeman, R., Miller, D., & Newcomer, L. (2015). Integration of academic and behavioral MTSS at the district level using implementation science. *Learning Disabilities--A Contemporary Journal*, 13(1), 59–72.
- Fuchs, S. L., Bucka, N., Clarke, B., Dougherty, B., Jordan, N. C., Karp, K. S., & Woodward, J. (2021). *Assisting Students Struggling with Mathematics: Response to Intervention (RtI) for Elementary and Middle Schools*. WWC 2021006. What Works Clearinghouse.
- Fuchs, D., & Fuchs, L. S. (2006). Introduction to Response to Intervention: What, why and how valid is it? *Reading Research Quarterly*, 41, 93–99. <https://doi.org/10.1598/RRQ.41.1.4>
- Fuchs, L. S., Fuchs, D., Craddock, C., Hollenbeck, K. H., Hamlett, C. L., & Schatschneider, C. (2008). Effects of small-group tutoring with and without validated classroom instruction on at-risk students' math problem solving: Are two tiers of prevention better than one? *Journal of Educational Psychology*, 100, 491–509. <https://doi.org/10.1037/0022-0663.100.3.491>
- Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). *Assisting Students Struggling with Mathematics: Response to Intervention (RtI) for Elementary and Middle Schools*. NCEE 2009-4060. What Works Clearinghouse.
- Hosp, M. K., Hosp, J. L., & Howell, K. W. (2007). *The ABCs of CBM: A practical guide to curriculum-based measurement*. (1st ed., pp. 1–172). New York: Guilford Press.
- Hunt, J. (2014). Effects of a supplemental intervention focused in equivalency concepts for students with varying abilities. *Remedial and Special Education*, 35, 135–144. <https://doi.org/10.1177/0741932513507780>
- Jitendra, A. (2007). *Solving math word problems: Teaching students with learning disabilities using schema-based instruction*. Pro-Ed.
- Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case designs technical documentation*. What Works Clearinghouse.
- Kratochwill, T. R., & Levin, J. R. (Eds.). (2014). *Visual analysis of single-case intervention research: Conceptual and methodological issues*. In T. R. Kratochwill & J. R. Levin (Eds.), *School Psychology Series. Single-Case Intervention Research: Methodological and statistical advances* (p. 91–125). American Psychological Association. <https://doi.org/10.1037/14376-004>
- Manolov, R., & Onghena, P. (2018). Analyzing data from single-case alternating treatments designs. *Psychological Methods*, 23(3), 480–504. <https://doi.org/10.1037/met0000133>
- Mason, E. N., Benz, S. A., Lembke, E. S., Burns, M. K., & Powell, S. R. (2019). From professional development to implementation: A district's experience implementing mathematics tiered systems of support. *Learning Disabilities Research & Practice*, 34(4), 207–214. <https://doi.org/10.1111/ldrp.12206>
- Massachusetts Department of Elementary and Secondary Education (2019). *School and District Profiles*. <https://profiles.doe.mass.edu/statereport/classsizebyraceethnicity.aspx>
- Mellard, D., McKnight, M., & Jordan, J. (2010). RTI tier structures and instructional intensity. *Learning Disabilities Research & Practice*, 25, 21–225. <https://doi.org/10.1111/j.1540-5826.2010.00319.x>
- National Center for Education Statistics (NCES) (2018). *The Nation's report card: 2017 Mathematics & Reading Assessments*. Institute of Education Sciences, U.S. Department of Education.
- Poncy, B. C., & Duhon, G. J. (2015). Measures & Interventions for Numeracy Development (MIND): Skill Remediation. *Intervention Manual*. https://b319dbee-747c-4d18-b9e4-407f0fd0b5b0.files.usr.com/ugd/7a057f_3c97bd2a7d674c22a9bb4e79a06ce7a3.pdf
- Poncy, B. C., Skinner, C. H., & Axtell, P. K. (2010). An investigation of detect, practice, and repair to remedy math-fact deficits in a group of third-grade students. *Psychology in the Schools*, 47, 342–352. <https://doi.org/10.1002/pits.20474>
- Poncy, B. C., Skinner, C. H., & Jaspers, K. E. (2007). Evaluating and comparing interventions designed to enhance math fact accuracy and fluency: Cover, copy, and compare versus taped problems. *Journal of Behavioral Education*, 16(1), 27–37. <https://doi.org/10.1007/s10864-006-9025-7>
- Schutte, G. M., Duhon, G. J., Solomon, B. G., Poncy, B. C., Moore, K., & Story, B. (2015). A comparative analysis of massed vs. distributed practice on basic math fact fluency growth rates. *Journal of School Psychology*, 53(2), 149–159.
- Shapiro, E. S. (2010). *Academic skills problems: Direct assessment and intervention*. Guilford Press.
- Shinn, M. R. (1989). *Curriculum-based measurement: Assessing special children*. Guilford Press.

- Silva, M., Coddling, R. S., Collier-Meek, M. A., & DeFouw, E. R. (2020). Acceptability assessment of school psychology interventions from 2005–2017. *Psychology in the Schools, 57*, 62–77. <https://doi.org/10.1002/pits.22306>
- Solomon, B. G., Poncy, B. C., Battista, C., & Campaña, K. V. (2020). A review of common rates of improvement when implementing whole-number operation math interventions. *School Psychology, 35*(5), 353–362. <https://doi.org/10.1037/spq0000360>
- Spectrum K12, American Association of School Administrators, Council of Administrators of Special Education, National Association of State Directors of Special Education, & State Title 1 Directors (2010). Response to intervention (RTI) adoption survey. http://sss.usf.edu/resources/presentations/2010/fasp_summer_inst2010/Resource_SLD/RTI/2010RTIAdoptionSurveyReport.pdf
- VanDerHeyden, A. M., & Burns, M. K. (2008). Examination of the utility of various measures of mathematics proficiency. *Assessment for Effective Intervention, 33*, 215–224. <https://doi.org/10.1177/1534508407313482>
- VanDerHeyden, A. M., McLaughlin, T., Algina, J., & Snyder, P. (2012). Randomized evaluation of a supplemental grade-wide mathematics intervention. *American Education Research Journal, 49*, 1251–1284. <https://doi.org/10.3102/0002831212462736>
- Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews, 13*, 70–77. <https://doi.org/10.1002/mrdd>
- What Works Clearinghouse. (2020). What Works Clearinghouse Standards Handbook, Version 4.1. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. <https://ies.ed.gov/ncee/wwc/handbooks>
- Yeaton, W., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology, 49*, 156–216. <https://doi.org/10.1037/0022-006X.49.2.156>
- Yoder, P. J., & Woynaroski, T. (2015). How to study the influence of intensity of treatment on generalized skill and knowledge acquisition in students with disabilities. *Journal of Behavioral Education, 25*, 152–166. <https://doi.org/10.1007/s10864-014-9216-6>

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