Hammill Institute on Disabilities

Remedial and Special Education
1–11

DOI: 10.1177/0741932520942954
rase.sagepub.com
SAGE

Effects of Formative Assessment Strategies on the Fractions Computation Skills of Students With Disabilities

Brian A. Bottge, EdD[1], Xin Ma, PhD[1], Linda J. Gassaway, EdD[1], Megan Jones, PhD[1], and Meg Gravil, PhD[2]

Abstract

Learning to compute with fractions is a major challenge for many students and especially for students with disabilities (SWD). Phase 1 of this study employed a randomized pretest–posttest comparison design to test the effects of two versions of formative assessment combined with an instructional program called Fractions at Work. In one condition, teachers used technology-assisted prompts to assess student performance and remediate errors. In the comparison condition, teachers gave students the same items for assessing progress but used their own methods of reteaching. Results indicated no difference between the two methods. However, pretest-to-posttest gain scores were significantly higher on all three measures regardless of type of formative assessment, and students maintained much of what they had learned. Phase 2 examined issues related to instructional dosage. Students who received additional weeks of instruction scored significantly higher than students who went back to their business-as-usual curriculum.

Keywords

instruction, mathematics, assessment, formal/informal, quantitative, research methodology

Results of the 1996 National Assessment of Educational Progress (NAEP) revealed that almost two thirds (65%) of eighth-grade students could not order, from smallest to largest, five fractions less than one (Silver & Kenney, 2000). More than 20 years later, large numbers of students continue to have difficulty computing with fractions. The most recent NAEP administrations (2011, 2013, 2015, 2017, and 2019) indicated that at eighth grade 64% to 69% of students with disabilities (SWD) and 21% to 25% of students without disabilities scored below Basic (McFarland et al., 2019 Mathematics Assessment). According to NAEP, students who score at the Basic performance level have only partial knowledge and skills for that grade, which suggests students who score *below* Basic have almost no ability to compute fractions.

Why are fractions so difficult? Research suggests that students may over-generalize whole-number properties and apply them to rational numbers (Behr et al., 1983, 1992; Radatz, 1979). For example, comparing two fractions with the same numerators, students show their *whole number bias* (e.g., Ni & Zhou, 2005; Vamvakoussi & Vosniadou, 2010) when they think that the fraction with the larger digit in the denominator is the larger unit. Students also make the mistake of adding or subtracting numerators and denominators as if they were whole numbers (e.g., 5/8 + 1/4 = 6/12; 3/4 – 1/2 = 2/2). Many students, even those without disabilities, cannot judge the magnitude of fractions, visualize the part-whole relationships they represent, or understand the concept of equivalency (Malone & Fuchs, 2017; Siegler et al., 2011). To fix their computation difficulties, students need to develop a basic conceptual understanding of rational numbers so the reasons for the procedural steps make sense. If not learned together, it is likely students will find fractions confusing and continue to make bizarre computation errors (Siegler & Pyke, 2013).

The Promise of Formative Assessment

In response to calls for improving student performance in key areas such as rational numbers, the Council of Chief State School Officers (CCSSO; Brookhart & Lazarus, 2017) has urged teachers to use Formative Assessment (FA). The CCSSO (2018) defined FA as a planned and

[1]University of Kentucky, Lexington, KY, USA
[2]University of Louisville, KY, USA

Corresponding Author:
Brian A. Bottge, University of Kentucky, 222 Taylor Education Building, Lexington, KY 40506, USA.
Email: brian.bottge@uky.edu

ongoing method of instruction designed to help teachers improve their students' academic performance. FA accomplishes this by clarifying learner goals and setting the criteria for success, providing a window into student thinking, encouraging students to improve their own self-assessment, and guiding teachers toward more effective teaching.

Much of what authors describe as FA practices is not new to special educators. Fundamental to special education is the promise that teachers will use strategies to monitor, assess, and revise their instruction based on the individual needs of their students. Progress monitoring has a rich history of practice in the form of Curriculum-Based Measurement (CBM; Deno, 1985) and Response to Intervention (RTI; D. Fuchs et al., 2012). More than 30 years ago, a meta-analysis of FA procedures yielded an average effect size of .70, suggesting they can make important improvements in student performance (L. S. Fuchs & Fuchs, 1986).

Despite enthusiasm over its benefits, obstacles for using FA remain. Some of the more serious concerns relate to the quality of the periodic measures, the ability of teachers to use the test information to inform remediation activities, and the motivation of students to be involved in the process (Phelan et al., 2011). Perhaps the most critical of these relates to questions concerning teacher preparedness for using the assessment data to inform remedial activities, especially when the subject area is complex as is the case with fractions computation. Unless teachers have the necessary background to understand why students make the errors and the pedagogical knowledge to do something about it, FA is not likely to result in higher achievement.

For several years our research teams have designed and tested units of instruction called *Fractions at Work-Basic* (FAW-B) to help teachers do a better job of teaching rational number concepts and fractions computation skills to students across a range of skill levels. We developed FAW-B out of necessity because students had difficulty computing the correct answers to our problem-solving units that required computation with whole numbers and fractions. FAW-B was successful in boosting students' computation, overall, but many students continued to make common errors on the posttests. Based on those test results and frequent classroom observations, we concluded that one cause of the problem had to do with teachers who did not make effective use of the skill check items found at the end of each lesson. We selected the items to assess specific skill areas teachers had taught students in the preceding instructional unit. Some teachers used them to identify the misconceptions that led to computation errors but they typically did not take the time to adequately reteach the concept.

Research Questions

To address these findings, we revised FAW (FAW-R) by adding instructional supports to the assessment items that were part of the original version of FAW-B. We hoped that FAW-R would provide teachers with more ideas for addressing conceptual and procedural errors of students as they moved through the FAW-B units. Our primary purpose of this study was to compare the effectiveness of the newly revised FAW-R with that of the previous version. We also wondered if students who learned with the new FAW-R could maintain their skills and what effect would extending instructional time have on performance. Specifically, we designed the study to answer the following questions:

1. **Research Question 1 (RQ1):** What are the differential effects, if any, of FAW-B and FAW-R on the fraction computation skills of SWD?
2. **Research Question 2 (RQ2):** To what extent do SWD maintain their fractions computation skills after instruction with FAW-R?
3. **Research Question 3 (RQ3):** What effect, if any, does lengthening instructional time have on the fraction computation skills of SWD?

Method

Participants

Table 1 displays demographic background of the 19 teachers and 110 students from the 11 schools who participated in the study. Students, parents, and teachers signed consent forms approved by the University Institutional Review Board. Due to the severity of students' math disabilities, the Admissions and Release Committees at each school placed the students in special education self-contained resource rooms where they received intensive, small-group instruction. The average pretest raw score on the standardized Mathematics Computation subtest of the *Iowa Tests of Basic Skills* (ITBS; Form C, Level 12; University of Iowa, 2008) of the SWD in this study was 9 (30% correct) compared with the raw score of 19 (64% correct) for the overall normed sample.

Most students were receiving special education services for specific learning disabilities and mild mental disabilities. The other students had been diagnosed with autism spectrum disorder, emotional and behavioral disorder, or other health impaired. For some students English was their second language. The FAW-B and FAW-R intervention groups were comparable in gender, ethnicity, and disability type (see Table 1). The class size of the 18 FAW-R resource rooms (M = 4.32, SD = 2.06) and the 8 FAW-B resource rooms (M = 3.44, SD = 1.91) did not differ. Classes met 5 days a week: 14 for 50 to 60 min, 8 for 45 to 49 min, and 4 for 61 to 67 min.

Research Design

Figure 1 indicates elements of the research design and the time spent in each condition. We randomly assigned schools

Table 1. Teacher and Student Characteristics.

Participant	Sequence A[a]	Sequence B[b]	χ^2	*t*	*p*
Teacher					
Gender			0.61 (1)		.44
Male	3	1			
Female	8	7			
Ethnicity			4.23 (3)		.24
Caucasian	9	6			
Asian	0	1			
African American	2	0			
Biracial and other	0	1			
Highest degree earned			0.06 (1)		.81
BA, BS	1	1			
MA, MS	10	7			
Years teaching special education				.16 (17)	.88
M	9.64	10.13			
Median	8.00	10.50			
SD	6.98	5.51			
Range	22.00	18.00			
Student					
Gender			0.19 (1)		.66
Boys	50	21			
Girls	29	10			
Grade			8.15 (5)		.15
6	28	8			
7	26	11			
8	25	5			
11	0	1			
12	0	6			
Ethnicity			2.41 (4)		.66
Caucasian	45	19			
Asian	1	0			
Latino	11	3			
African American	19	6			
Biracial and other	3	3			
Disability/service area			8.15 (5)		.15
MMD	22	6			
OHI	19	11			
EBD	2	4			
Autism	13	2			
SLD	22	8			
FMD	1	0			
Free/subsidized lunch			0.33 (1)		.56
No	15	3			
Yes	61	8			

Note. Values in parentheses represent degree of freedom. MMD = mild mental disability; OHI = other health impaired; EBD = emotional/behavioral disability; SLD = specific learning disability; FMD = functional mental disability. Free/subsidized lunch has missing data on 23 students.
[a]*n* = 11 for teachers, 79 for students. [b] *n* = 8 for teachers, 31 for students.

to Sequence A (SeqA) or Sequence B (SeqB). We made the assignment by school to prevent possible contamination of intervention procedures within schools. In some schools, more than one teacher participated. FAW-R classrooms outnumbered FAW-B classrooms because the random selection included more schools with multiple teachers.

Teachers in each sequence administered the same two tests at three time points (O_1, O_2, and O_3). Phase 1 study (O_1–O_2) employed a pretest–posttest comparison design to assess the differential effects, if any, between FAW-B and FAW-*R* (RQ1). In Phase 2, we explored two issues related to instructional dosage: Skills maintenance following

		Phase 1		Phase 2	
SeqA:	O_1	$X_{FAW\text{-}R}$ (53 days) (2,802 minutes)	O_2	X_{BAU} (20 days) (1,010 minutes)	O_3
SeqB:	O_1	$X_{FAW\text{-}B}$ (41 days) (2,125 minutes)	O_2	$X_{FAW\text{-}R}$ (26 days) (1,320 minutes)	O_3

Figure 1. Research design and amount of instructional time in each phase.

instruction with FAW-R (RQ2) and the effects of additional time on performance (RQ3). We note here that it took longer to teach FAW-R than FAW-B because of the expanded FA activities.

Instructional Procedures

During Phase 1, teachers in FAW-B and FAW-R taught the same seven units in the core instructional program. The units addressed the fraction subconstructs of *length* and *set* and emphasized three concepts that students in previous studies found most confusing: part-whole partitioning, ratio, and measurement. We supplied teachers with daily lesson plans.

The introductory unit focused on basic but important concepts such as why it is important to learn fractions, how fractions relate to whole numbers, and the roles of numerators and denominators. After viewing the first instructional video, students worked on two hands-on activities. The first was called the *Big Inch Ruler*. The teacher handed out fractions labels that students placed on their large inch ruler, which was also projected onto a white board. Students discussed where to place the labels onto the inch. This was an informal activity to assess students' prior knowledge of fractions. In the second activity, students cut out fraction strips labeled with multiples (e.g., 2, 4, 6, 8 and 3, 6, 9, 12), which the students would use as a tool in later lessons for finding equivalent fractions.

Most lessons also included opportunities for learning with manipulatives. For example, in one of the lessons the teacher gave students four long strips of paper and told them to imagine each strip was a candy bar. The teachers asked students to pretend they were sharing the candy bar with two people. The students folded the strip in half and labeled the crease 1/2 and one end 2/2. The teacher then asked students to fold the "candy bar" twice to make four "pieces." This was repeated to represent eight and 16 people sharing the candy. The teacher then told the class to lay the strips side-by-side with the creases from the folds lining up across the strips. The teacher followed up this activity with a discussion of questions such as: *What is another name for ¾?* Each student was expected to show it with fraction strips. Teachers referred to this lesson often to remind students the function of denominators and numerators. The next lesson showed a large inch on an interactive tape measure to help students understand that the value of a fraction depends on the number of parts into which an inch is divided (i.e., denominator) and the number of these parts that are available (i.e., numerator).

In the lessons that followed, the software used animations to show that as the denominator number increases the size of the fraction decreases. The media also demonstrated how to add simple fractions with like and unlike denominators, add and subtract mixed numbers, and rename and simplify fractions. The lesson plans continued the practice of directing teachers and students back to the fraction strips to find equivalent fractions and to answer questions like these: *If you have 3/4 of a stick of gum and your friend has 1/2 of a stick, who has more? How much more does she have?*

Two FA Conditions

Each intervention condition included either the basic FAW-B or the revised version (FAW-R). As noted previously, teachers taught all students the same instructional lessons (FAW). The two conditions differed by the structure of the FAs and how teachers delivered the corrective instruction. FAW-R included technology-based tools to help students (and teachers) recognize conceptual misunderstandings that led to computing errors. FAW-B made use of the same assessment items but left detection and reteaching methods up to the teacher.

Fractions at Work-Revised. We developed the new Fractions at Work-Revised (FAW-R) with three goals in mind. First, we wanted to give teachers an opportunity to identify common errors each student made so they could target their reteaching to fix those errors. Second, we anticipated students would be able to correct some of their own errors after working on the FAW-R materials thereby decreasing dependency on teacher prompting. Third, we hoped that the more explicit instructional activities would model for teachers how to reteach fraction computation skills consistent with the lesson objectives.

FAW-R consisted of 60 fractions computation practice problems and prompts organized into sets of 10 and distributed over six of the seven FAW-R units. We carefully selected each practice problem to match the learning objectives of the FAW-R unit, which students had just learned in the previous lesson. The research team studied each of the 60 assessment problems and decided the best way of prompting students' thinking. After the team members agreed on the content of the support, they discussed the level of support most appropriate for that item: For example, narration + think aloud, narration + think aloud + still

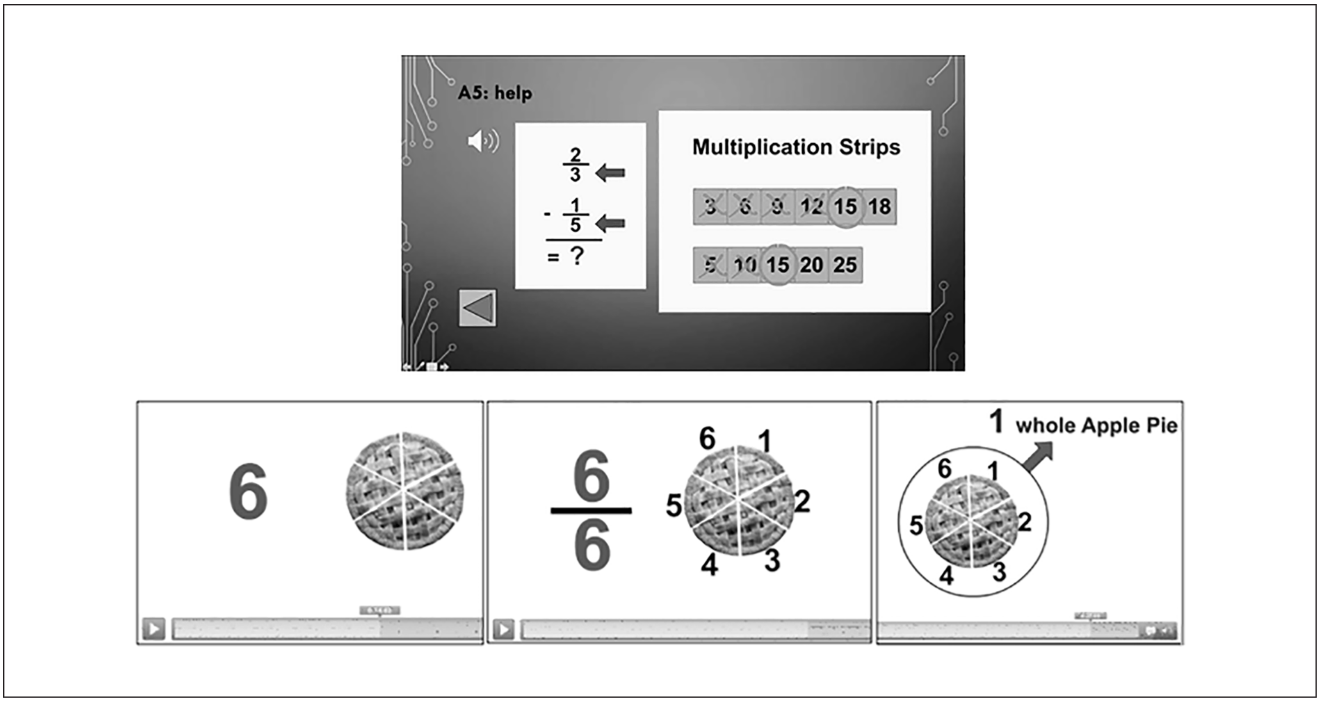

Figure 2. Panel A: FAW-R visual of common multiples for finding common denominators and Panel B. FAW-R animated visual showing functions of numerators and denominators.
Note. FAW-R = Fractions at Work–Revised.

visuals, and narration + think aloud + visual animation. The lesson plans called for teachers to display one-by-one each of the 10 practice problems on the projection screen.

About half (31) of the practice problems were narrated with a think aloud prompt. For example, one item asked students to solve this problem: 1/2 + 3/8 = . After giving students time to solve the problem on their own, the teacher played this support: *This is an addition problem with fractions that do not have common denominators. Can I multiply the 2 by some number to make 8 (the denominator in the other fraction)?* Students then chose one of the answers: A. 5/12, B. 7/8, C. 5/27, or D. 5/8.

We also added visual supports to 16 of the items such as: *Choose the correct answer: 2/3 – 1/5 = ?* The support screen (see Figure 2) showed two images during the audio clip: *The denominators are not common and 5 is not a multiple of 3. I need to find the smallest common multiple of 5 and 3. I can use my fraction strips* (see Panel A of Figure 2). After using the prompt to solve the problem, students made their choice from the following options: A. 1/2, B. 7/15, C. 1/5, or D. 7/30. Some items (13) included a short animation with voiceover (see Panel B, Figure 2): *If the numerator is the same as the denominator, I know the fraction is equal to 1. For example, if an apple pie is cut into 6 pieces, that would be the denominator, and if all 6 pieces remain, that would be the numerator. I have 6 over 6. So, I still have the whole apple pie. 6/6 is equal to 1.*

All computation problems were either addition or subtraction of simple or mixed number fractions. Most items asked students to solve the problem and to select the correct answer from four possible choices (one correct answer and three distractors). The distractors reflected the most common errors students had committed on similar problems in previous studies (e.g., identifying 1/4 as larger than 2/3; whole number bias; Suh et al., 2018). The other items were open-ended and required students to supply the correct answer.

Fractions at Work-Basic. Teachers who taught Fractions at Work-Basic (FAW-B) used the same lesson plans for teaching fractions concepts and procedures as those in the new FAW-R program but they did not have access to the supports included in FAW-R. Teachers showed the 10 computation items on the screen one-by-one while the students followed along on a worksheet. After teachers showed the problem and read it aloud, students worked the problem on their own and then selected their answer(s) from multiple choice items, or wrote their answer in the blank. After students finished a problem, teachers decided what to do with the information provided by the students' work product. The lesson plans suggested that teachers examine each student's work on the problems and consider how to correct misconceptions by reviewing relevant instructional activities in FAW. Teachers were to either provide this review on the same day or on the days that followed but prior to working on the new set of problems.

Professional Development

In preparation for the study, the project manager spent 2 hr at each school training participating teachers on specific procedures for teaching the lessons and administering the FAs. We chose the individualized one-on-one training format to address the specific instructional and assessment needs of each teacher. From our previous experiences working with teachers and technology, we knew that the individualized training format was necessary because the software had to be installed on school computers and then tested to ensure that it ran smoothly.

The content of the training sessions consisted of an introduction to the concept of FAW-B or FAW-R, detailed descriptions of each lesson, the purpose and function of hands-on manipulatives such as fraction strips, and the procedures for conducting the pretests, posttests, and FAs. A portion of the session was also devoted to procedures for obtaining Institutional Review Board (IRB) consent, and how to complete the daily instructional logs and lesson plan checklists that we required as part of implementation fidelity records.

Instrumentation

Teachers administered two tests on consecutive days immediately before and after each study phase. *The Fractions Computation Test* (FCT) is a 20-item (14 addition, 6 subtraction), 42-point researcher-developed measure that assessed students' ability to add and subtract simple fractions and mixed numbers with like and unlike denominators (Figure 3). The test also included four items that required adding three fractions. All but one of the items contained fractions found on a ruler but students did not have access to rulers nor could they use calculators during the testing sessions. Teachers asked students to show all their work and to simplify their answers. Students could earn one point for showing correct work and one point for the correct answer on 18 of the items. They could earn an additional point on two items with mixed numbers that required renaming prior to subtracting. Internal consistency estimates O_1, O_2, and O_3 were 0.75 (95% confidence interval [CI] = [0.68, 0.82]), .95 (95% CI = [0.94, 0.97], and 0.96 (95% CI = [0.94, 0.97]), respectively.

The second measure was a 30-item ITBS Computation (ITBSC) subtest of the ITBS (Form C, Level 12; University of Iowa, 2008) that measured computation (addition, subtraction, multiplication, and division) with whole numbers, fractions, and decimals. Students selected the correct answer from four choices. One constant choice was *N* indicating that the other three choices did not include the correct answer. Internal consistency estimates reported for the total subtest items for O_1, O_2, and O_3 were 0.71 (95% CI = [0.63, 0.78], 0.76 (95% CI = [0.69, 0.82]), and 0.75 (95% CI [0.68, 0.82]), respectively.

From the ITBSC we selected the eight items that tested students' ability to add and subtract fractions. We used these items as a group to provide another indicator of students' performance on fractions computation apart from the ITBSC whole-number problems.

Implementation Fidelity

Three graduate assistants and the project manager conducted a combined total of 124 whole-class period observations. Observers entered data directly into a computer-based collection tool that included space to record demographic information and levels of student engagement, open areas for describing quality of implementation, and a rating scale for reporting the degree of fidelity. To supplement the in-class observations, teachers completed daily Instructional Logs and Lesson Checklists. In their Instructional Logs, teachers indicated which study-related activity they taught that day. If they did not teach the study curriculum, they noted the reason (e.g., field trip, school-wide assembly, teacher illness). This information enabled the researchers to calculate how much instructional time teachers devoted to each lesson. In addition, teachers recorded class attendance each day, completed a Lesson Checklist for each lesson of the unit, and rated themselves daily on how closely they followed the lesson plans. Classroom visitations by a second observer were minimal (12) because of the locations of the study sites. We found only minor inconsistencies between these information sources.

Results

Pretest scores (O_1) indicated no significant difference between SeqA and SeqB groups on the FCT ($t = –0.17$, $df = 107$, $p = .86$) and the ITBS ($t = 0.94$, $df = 108$, $p = .30$). We used a two-level hierarchical linear model (HLM; Raudenbush & Bryk, 2004) with students nested within teachers to examine student performance on each of the math outcomes. We chose not to use the three-level model (with students nested within teachers nested within schools) because the sample was very small (one to three teachers) at the school level. The two-level HLM "returned" some of the variance from the school level to the teacher level, thus allowing for a better modeling of teacher effects. The two-level HLM had scores at the second and third observations (O_2 and O_3) as the outcomes and controlled for five background variables at the student level (gender, race-ethnicity, and disability status with three dummy variables), together with the score from the first observation (i.e., the pretest score).

Table 2. Simplified Two-Level HLM Models Comparing FAW-R to FAW-B in a Pretest and Posttest Experimental Design (Observation 2 vs. Observation 1).

Measure	Effects	*SE*	R^2
ITBS overall computation	1.06	0.98	.20
ITBS fraction computation	0.04	0.58	.16
Fraction computation test	–2.43	2.70	.25

Note. R2 = proportion of variance explained. The two-level HLM model includes Observation 1 scores and student characteristics of gender, race, and disabilities at the student level as well as a dummy variable denoting *FAW-R* vs. *FAW-B* at the teacher level. Because the HLM program does not allow any missing value on any outcome variable, data analysis is based on a different number of students for a given cognitive measure who have valid Observation 2 scores. Estimates on ITBS overall computation and ITBS fraction computation are based on 109 students nested within 19 teachers, while estimates on fraction computation test are based on 108 students nested within 19 teachers. HLM = hierarchical linear model; FAW-B = Fractions at Work–Basic; FAW-R = Fractions at Work–Revised; ITBS = Iowa Tests of Basic Skills.
$^*p < .05$.

Table 3. Simplified Two-Level HLM Models Estimating Value-Added Benefit of Additional Instruction Time (Observation 3 vs. Observation 2).

Measure	Effects	*SE*	*ES*	R^2
ITBS overall computation	1.45*	0.64	.29	.34
ITBS fraction computation	0.59*	0.26	.37	.39
Fraction computation test	6.86*	2.57	.70	.49

Note. R^2 = proportion of variance explained. The two-level HLM model includes Observation 2 scores and student characteristics of gender, race, and disabilities at the student level as well as a dummy variable denoting additional instruction time (vs. traditional math curriculum) at the teacher level. Because the HLM program does not allow any missing value on any outcome variable, data analysis is based on a different number of students for a given cognitive measure who have valid Observation 3 scores. Estimates on ITBS overall computation and ITBS fraction computation are based on 109 students nested within 19 teachers, while estimates on fraction computation test are based on 108 students nested within 19 teachers. ES = effect size; HLM = hierarchical linear model; ITBS = Iowa Tests of Basic Skills.
$^*p < .05$.

We conducted several HLMs to address the research questions. Take as an example the issue of differential effects between FAW+B and FAW+R on the ITBS Overall Computation with ITBS_O_1 as the first observation score and ITBS_O_2 as the second observation score, the student-level model was

$$ITBS\_O_{ij} = \beta_{0j} + \beta_{1j} ITBS\_O1_{ij} + \sum_{n=1}^{5} \beta_{(n+1)j} SC_{nij} + \varepsilon_{ij}$$

where ITBS_$O2_{ij}$ was the computation score at the second observation for student i with teacher j, ITBS_$O1_{ij}$ was the computation score at the first observation for the same student, and ε_{ij} was an error term unique to each student, $\varepsilon_{ij} \sim N(0, \sigma^2)$. Student (background) characteristics, SC_{nij} (n = 1, 2, 3, 4, 5), were used as the control variables. The average computation score of students at the second observation was captured in β_{0j} for teacher j adjusted for student characteristics and computation score at the first observation for that teacher. Thus, β_{0j} became the outcome at the teacher level. All slopes or coefficients from the student level were fixed without variance at the teacher level.

With β_{0j} as the outcome, the teacher-level model included the indicator of experiment condition. Specifically, the teacher-level model was

$$\beta_{0j} = \gamma_{00} + \gamma_{01} C_j + u_{0j}$$

where γ_{00} was the average computation score at the second observation and u_{0j} was an error term unique to each teacher, $u_{0j} \sim N(0, \tau)$. The dummy variable C_j used FAW-B as the baseline against which FAW-R was compared. Results from this HLM were used to create Table 2. A similar HLM focusing on O_2 and O_3 generated results for Table 3. Overall, these HLMs employed a full information maximum likelihood estimation method, using all available data except those missing on a certain outcome. All statistical significance tests were performed at the alpha level of .05. These HLMs also yielded ICC (intraclass correlation) which, based on the partition of variance, measured the proportion of variance in a certain outcome for which teachers were responsible. Proportion of variance explained, referred to as R^2, was used to evaluate the performance of HLMs. Finally, we used Hedges' g (1981) as the effect size (*ES*) measure whenever appropriate.

Overall Findings

Table 4 shows means and standard deviations of the ITBSC subtest, the eight ITBS fractions items, and the FCT at three time points for students by instructional condition. The ICCs indicated sizable teacher effects, which suggests that teachers were differentially effective in teaching fractions to the SWD. Several trends emerged from the descriptive data. First, most scores increased from pretest to posttest, which suggests that the FA methods were effective. Second, the size of SDs on the FCT posttests suggests a wide variation in the performance between students who profited from instruction and those who did not.

Table 2 compares the effects of SeqA and SeqB for time points O1 and O2 for each measure. No statistically significant differences between the two experimental conditions were found. However, additional analyses suggest that scores from O1 to O2 of students pooled without regard to intervention method showed significant gains on all three measures (see Appendix A). Effect sizes were .57 on the ITBS Overall Computation, .25 on the ITBS Fraction Computation, and 1.32 on the FCT.

Table 4. Descriptive Statistics for Students ($n = 110$) on ITBS Overall Computation, ITBS Fraction Computation, and Fraction Computation Test by SeqA and SeqB.

Measure	ICC	SeqA		SeqB	
		M	*SD*	*M*	*SD*
ITBS computation (30)					
Observation 1		8.65	4.06	9.58	4.60
Observation 2	.08	11.86	4.59	11.48	5.23
Observation 3	.06	11.65	4.94	12.50	4.02
ITBS fraction subset (8)					
Observation 1		3.11	1.53	2.71	1.51
Observation 2	.27	4.62	1.65	4.58	2.32
Observation 3	.14	4.44	1.58	4.96	1.73
Fraction computation (42)					
Observation 1		2.85	3.37	2.71	4.37
Observation 2	.17	10.39	11.10	13.35	12.60
Observation 3	.22	8.32	9.73	16.45	13.86

Note. Values in parentheses after the outcome measure represent the number of points possible. ITBS = Iowa Tests of Basic Skills; ICC = intra-class correlation.

Table 3 indicates the effects of adding additional instructional time to SeqB for time points O2 to O3. The added time significantly improved students' test scores on the FCT, the ITBS, and the eight ITBS fractions items compared with students who went back to their business-as-usual math instruction. Effect sizes ranged from .29 on the full ITBSC test to .70. on the FCT. Performance on the eight fractions computation items of the ITBS also showed an ES of .37. Additional analyses suggest that students in SeqA who returned to their Business-As-Usual lessons maintained their level of performance from O2 to O3 (see Appendix B).

Discussion

The purpose of this study was to test the effects of two sequences of instruction on the fractions computation skills of SWD measured across two phases. In the first phase we employed a randomized pretest–posttest design to compare two FA strategies each combined with an instructional program called the basic version Fractions at Work (FAW-B). The new FAW-R provided a set of technology-based tools for helping teachers to assess and remediate student work on fractions computation problems. The comparative condition (FAW-B) consisted of the same units of instruction and FA assessment items but did not contain the additional supports included in FAW-R.

Results were mixed. First, test scores pooled across the two strategies showed significant improvement on each of the three measures, which supports the use of the basic instructional program curriculum we had used in prior studies (Bottge et al., 2014). Second, we thought that the FAW-R would have a greater impact on computation skills than the basic version. However, results indicated no significant difference between FAW-B and FAW-R. In fact, students in FAW-B scored about three points higher on the posttest than students in FAW-R.

What could account for this finding? We can only speculate based on notes from classroom observations and our experiences in similar studies. First, although teachers in the FAW-B group did not have access to the technology-based enhancements in FAW-R, they used their own informal ways of monitoring each student's progress. The advantage of small class size in the resource rooms afforded teachers opportunities to assess their students' thinking and make appropriate adjustments to their instruction. For example, the observer noted that one teacher sat at a table with her students during each observation. The five students in this class were frequently engaged and volunteered answers/ideas with the group. The teacher guided students step-by-step through each problem with probing questions. Students frequently asked each other for help.

Second, teachers tended to use a mix of procedures in carrying out the daily lesson plans. We observed them incorporating their own ways of gauging student progress. Some teachers treated each set of FA problems as quizzes requiring students to work individually and in silence. Teachers played the question stem, followed by the "help" or audio/video/animation support. In one classroom, the teacher worked out the first five problems on the whiteboard while students followed along. She then asked students to complete the remaining five problems individually. The most important take-away from Phase 1 was the improvement students made from pretest to posttest regardless of which FA strategy teachers used.

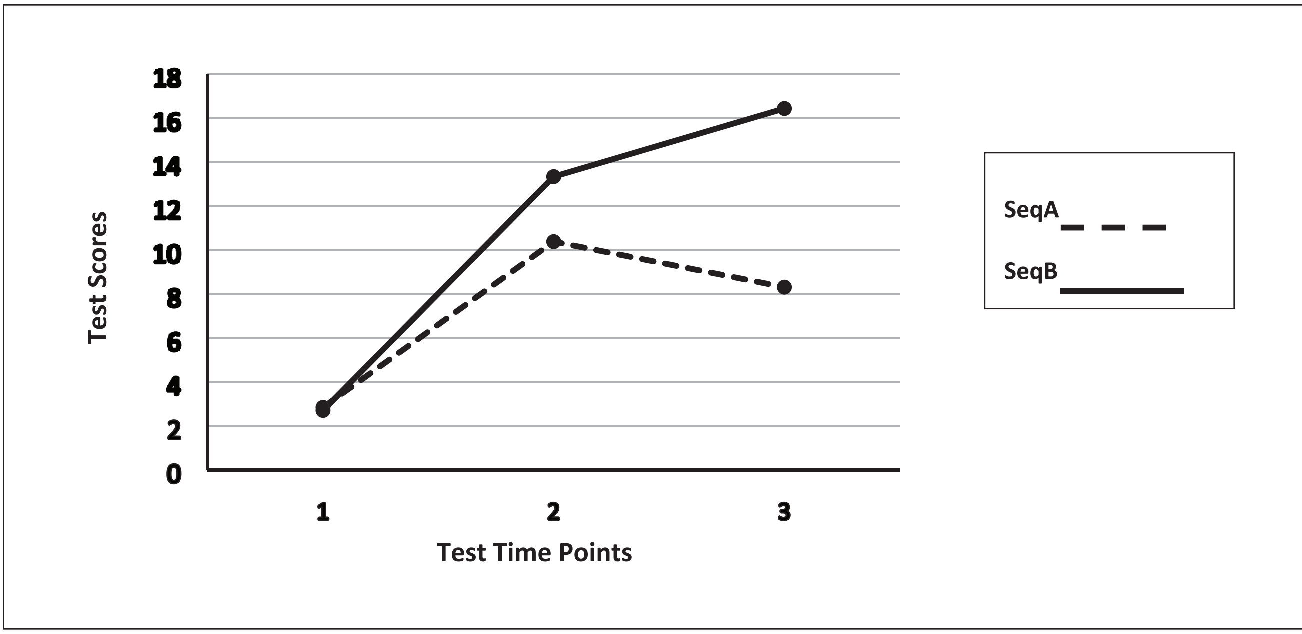

Figure 3. Mean fractions computation test scores SeqA and SeqB. Fractions Computation Test (Bottge).

In the second phase of the study, we examined issues related to instructional dosage. Students in SeqA who had been in the FAW-R condition for 10 weeks returned to their usual curriculum. Tests administered after the 4 weeks of business-as-usual instruction, which included math objectives other than rational numbers, indicated no significant decline in students' performance from the previous test administration. Meanwhile, teachers in SeqB who had been in the FAW-B condition for 8 weeks provided an additional 5-week dose of fractions instruction by reviewing and reteaching parts of the FAW-R units. The added time proved important because scores were significantly higher on all three measures. Observers noted behavioral changes as well. Some students became more confident in their ability and wanted to solve the problems on their own. We observed several instances when students wanted less help from teachers: "No, stop! I want to do it myself," "Let me try it," "Boy, I can figure this out," and "Please don't tell me! I got this! Why am I yelling?"

Finally, several of the observers wrote about how interesting it was to hear the SWD discuss fractions concepts with math vocabulary. For example one teacher asked *What did you do on this, J? What'd you do with 3?* to which J answered *I found the least common multiple*. Most students had not used this terminology prior to participating in this study.

The question about the time it takes to make a meaningful progress in learning difficult concepts such as rational numbers is open to question. Of the 21 studies reviewed in a recent meta-analysis of studies aimed at improving the fractions computation of SWD (Ennis & Losinski, 2019), about half of them spanned less than 25 days. This study along with our recent ones suggest that we should allow more than double that number of instructional days to make any kind of improvement in this difficult set of concepts and procedures. Allowing students sufficient time to learn difficult subject matter is one of the key recommendations in the Institute of Education Sciences Practice Guide, *Organizing Instruction and Study to Improve Student Learning* (Pashler et al., 2007).

Learning fraction concepts and procedures requires both domain-specific knowledge (e.g., judging numerical magnitudes) and more general cognition attributes (e.g., working memory, attention; Hansen et al., 2015; Kong, 2008; Siegler et al., 2011). Knowing how to weave these two forms of knowledge into an effective curriculum is especially challenging with rational numbers.

Study Limitations

We acknowledge two main limitations of our study. First, the design lacked a true control group. Instead, we compared the effects of two FA formats combined with a common set of instructional units. Adding a control group would have enabled us to sort out and identify the true effects of each intervention method.

Second, the length of staff development was only 2 hr, which is far less contact than the two 8-hr days we have usually devoted to training in our previous research. We thought the individualized training would help us to describe much of the detail for teaching FAW-B and FAW-R and using the FA tools. Although we found that teachers followed the lesson plans closely, we would advise lengthening staff development to at least a half day to allow for deeper coverage of the FAW instructional units and expanded discussion of FA formats.

Conclusion

Our findings support previous research that suggests FAW is an effective instructional program for improving the fractions computation skills of SWD. An important feature of FAW are the frequent skill checks spaced throughout the units. The criteria for judging the quality of these formative methods depend on their effects on the student. As demonstrated in this study, teachers can be effective using FAW with a prescribed set of technology-based tools, their own ways of checking for understanding with guided questioning and modeling, or in combination. Teachers should allow students enough time to acquire deep conceptual understanding of rational numbers so the procedural skills make sense and are maintained.

Appendix A

Table A1. Gain Scores All Students on Overall ITBS Computation, ITBS Fractions, and Computation Test (Observation 2 vs. Observation 1).

Measures	Range of score	Gain score	*SE*
ITBS overall computation	(0, 30)	2.74*	0.57
ITBS fraction computation	(0, 8)	1.63*	0.25
Fraction computation test	(0, 42)	8.71*	1.32

Note. The two-level HLM model uses difference scores between Observation 2 and Observation 1 as the outcome variable. Because the HLM program does not allow any missing value on any outcome variable, data analysis is based on a different number of students for a given cognitive measure who have valid difference scores. Estimates on ITBS overall computation and ITBS fraction computation are based on 109 students nested within 19 teachers, while estimates on fraction computation test are based on 107 students nested within 19 teachers. ITBS = Iowa Tests of Basic Skills; HLM = hierarchical linear model.
$^*p < .05$.

Appendix B

Table B1. Maintenance Scores on Overall ITBS Computation, ITBS Fractions Computation, and Fraction Computation Test (Observation 3 vs. Observation 2).

Measures	Range of score	Maintenance score	*SE*
ITBS overall computation	(0, 30)	–0.08	0.45
ITBS fraction computation	(0, 8)	–0.16	0.18
Fraction computation test	(0, 42)	–2.67	1.28

Note. The two-level HLM model uses difference scores between Observation 3 and Observation 2 as the outcome variable. Because the HLM program does not allow any missing value on any outcome variable, data analysis is based on a different number of students for a given cognitive measure who have valid difference scores. Estimates on ITBS overall computation and ITBS fraction computation are based on 77 students nested within 11 teachers, while estimates on fraction computation test are based on 76 students nested within 11 teachers. ITBS = Iowa Tests of Basic Skills; HLM = hierarchical linear model.
$^*p < .05$.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: The research reported in this article was supported by a grant from the U.S. Department of Education, Institute of Education Sciences (R324A150035). Any opinions, findings, or conclusions are those of the authors and do not necessarily reflect the views of the supporting agency.

ORCID iD

Brian A. Bottge https://orcid.org/0000-0002-7543-9996

References

Behr, M. J., Harel, G., Post, T., & Lesh, R. (1992). Rational number, ratio, and proportion. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 296–333). Macmillan.

Behr, M. J., Lesh, R., Post, T. R., & Silver, E. A. (1983). Rational number concepts. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematical concepts and processes* (pp. 91–126). Academic Press.

Bottge, B. A., Ma, X., Gassaway, L., Toland, M., Butler, M., & Cho, S. J. (2014). Effects of blended instructional models on math performance. *Exceptional Children*, *80*, 423–437. https://doi.org/10.1177/0014402914527240

Brookhart, S., & Lazarus, S. (2017). *Formative assessment for students with disabilities. commissioned by the Council of Chief State School Officers State Collaboratives on Assessing Special Education Students and Formative Assessment*. https://ccsso.org/sites/default/files/2017-12/Formative_Assessment_for_Students_with_Disabilities.pdf

Council of Chief State School Officers. (2018). *Revising the definition of formative assessment*. https://ccsso.org/resource-library/revising-definition-formative-assessment

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, *52*, 219–232.

Ennis, R. P., & Losinski, M. (2019). Interventions to improve fraction skills for students with disabilities: A meta-analysis. *Exceptional Children*, *85*, 367–386.

Fuchs, D., Fuchs, L. S., & Compton, D. L. (2012). Smart RTI: A next-generation approach to multilevel prevention. *Exceptional Children*, *78*, 263–279.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, *53*, 199–208.

Hansen, N., Jordan, N. C., Fernandez, E., Siegler, R. S., Fuchs, L., Gersten, R., & Micklos, D. (2015). General and math-specific predictors of sixth-graders' knowledge of fractions. *Cognitive Development*, *35*, 34–49. https://doi.org/10.1016/j.cogdev.2015.02.001

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 106–128.

Kong, S. C. (2008). The development of a cognitive tool for teaching and learning fractions in the mathematics classroom: A design-based study. *Computers & Education*, *51*, 886–899.

Malone, A. S., & Fuchs, L. S. (2017). Error patterns in ordering fractions among at-risk fourth-grade students. *Journal of Learning Disabilities*, *50*(3), 337–352. https://doi.org/10.1177/0022219416629647

McFarland, J., Hussar, B., Zhang, J., Wang, X., Wang, K., Hein, S., Diliberti, M., Forrest Cataldi, E., Bullock Mann, F., & Barmer, A. (2019). *The condition of education 2019* (NCES 2019-144). National Center for Education Statistics, U.S. Department of Education. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2019144

Ni, Y., & Zhou, Y. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist*, *40*, 27–52. https://doi.org/10.1207/s15326985ep4001_3

Pashler, H., Bain, P., Bottge, B., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J. (2007). *Organizing instruction and study to improve student learning* (NCER 2007-2004). National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. http://ncer.ed.gov

Phelan, J., Choi, K., Vendlinski, T., Baker, E., & Herman, J. (2011). Differential improvement in student understanding of mathematical principles following formative assessment intervention. *The Journal of Educational Research*, *104*(5), 330–339. https://doi.org/10.1080/00220671.2010.484030

Radatz, H. (1979). Error analysis in mathematics education. *Journal for Research in Mathematics Education*, *10*(3), 163–172.

Raudenbush, S. W., & Bryk, A. S. (2004). *HLM 6 for Windows* [Computer software]. Scientific Software International.

Siegler, R., & Pyke, A. (2013). Developmental and individual differences in understanding of fractions. *Developmental Psychology*, *49*(10), 1994–2004.

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology*, *62*(4), 273–296. doi:https://doi.org/10.1016/j.cogpsych2011.03.001

Silver, E. A., & Kenney, P. A. (2000). *Results from the seventh mathematics assessment of the National Assessment of Educational Progress*. National Council of Teachers of Mathematics.

Suh, Y., Cho, S.-J., & Bottge, B. A. (2018). A multilevel longitudinal nested logit model for measuring changes in correct response and error types. *Applied Psychological Measurement*, *42*, 73–88. https://doi.org/10.1177/0146621617703182

University of Iowa. (2008). *The Iowa Tests of Basic Skills* [ITBS, Form C]. Riverside Publishing.

Vamvakoussi, X., & Vosniadou, S. (2010). How many decimals are there between two fractions? Aspects of secondary school students' understanding of rational numbers and their notation. *Cognition and Instruction*, *28*(2), 181–209.