ENHANCING TEACHER INSTRUCTION THROUGH EVIDENCE-BASED EDUCATIONAL TECHNOLOGY: EVALUATING TEACHER'S USE OF DIFFERENTIATED INSTRUCTION

by

HAILEY HINKLE

A DISSERTATION

Presented to the Department of Special Education and Clinical Sciences and the Graduate
School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

December 2020

DISSERTATION APPROVAL PAGE

Student: Hailey Hinkle

Title: Enhancing Teacher Instruction through Evidence-Based Educational Technology: Evaluating Teacher's Use of Data Based Decision Making and Differentiated Instruction

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of School Psychology degree in the Special Education and Clinical Sciences by:

Hank Fien	Advisor/Chairperson
Nancy Nelson	Core Member
John Seeley	Core Member
Elizabeth Stormshak	Institutional Representative

and

Kate Mondloch	Interim Vice Provost and Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded December 2020.

DISSERTATION ABSTRACT

Hailey Hinkle

Doctor of Philosophy

Department of Special Education and Clinical Sciences

December 2020

Title: Enhancing Teacher Instruction through Evidence-Based Educational Technology: Evaluating Teacher's Use of Data Based Decision Making and Differentiated Instruction

Response to intervention, direct instruction, and data-based decision-making are all effective teaching practices that can impact student math performance. When educational technology is implemented in classroom settings, it can address both the needs of teachers and students. However, there is little information detailing how educational technology could enhance or impeded teaching practices like data-based decision-making and differentiated instruction for students with math learning difficulties.

A single-case research design was used to evaluate the quantity and quality of data-based decision-making and differentiated core math instruction by 1st-grade classroom teachers to support students with math learning difficulties. Teacher pre-and post-test measures, classroom observations, and technology data were all used to evaluate how teachers used an enhanced teacher data dashboard to support student math achievement in their classroom.

Omnibus and individual Tau-U effect sizes, visual analysis, and descriptive data were used to discuss the various technological features teachers used to support student math achievement. Classroom observations of teacher models, academic feedback, and student practice opportunities were observed. Information was also gathered regarding the implementation of evidenced-based educational technology in real classroom settings. Results indicated it was easier for teachers to change their instruction or opportunities for student practice than it was for them to provide corrective or affirmative feedback. The two teacher participants who have previous experience implementing educational technology in their classroom showed a greater capacity to engage in data-based decision-making and differentiated instruction than the two novice teacher participants. Implementation results indicated that the level of intervention intensity prescribed by the research team may be too intimidating for teachers to successfully implement in their classrooms for 12 consecutive weeks. Lastly, this study hopes to serve as an example for researchers on how intervention efficacy and implementation can be studied simultaneously within the context of a single-case design to study the unique ways participants engage with evidenced-based educational technology and interventions.

CURRICULUM VITAE

NAME OF AUTHOR: Hailey Hinkle

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
University of Central Oklahoma, Edmond

DEGREES AWARDED:

Doctor of Philosophy, School Psychology, 2020, University of Oregon
Masters of Art, School Psychology, 2014, University of Central Oklahoma
Bachelor's of Arts, Psychology, 2011, University of Central Oklahoma

AREAS OF SPECIAL INTEREST:

Evidence-Based Behavioral and Academic Interventions
Single Case Research Designs
Implementation of Evidence-Based Interventions

PROFESSIONAL EXPERIENCE:

School Psychologist, Longview School District, 2020-Present]

School Psychologist Intern, North Clackamas School District, 2019-2020]

Research Assistant, Prevention Science Institute, 2019-2020]

Graduate Research Assistant, Center on Teaching and Learning, 2019-2020]

Graduate Research Assistant, Center on Teaching and Learning, 2018-2019]

Graduate Research Assistant, Behavioral Research and Teaching, 2016-2018]

Diagnostic Intern, University of Central Oklahoma Learning and Behavior Clinic, 2015]

Behavioral Supervisor, Applied Behavior Analysis of Okalahoma, 2014-2015]

ACKNOWLEDGMENTS

I was incredibly fortunate to be surrounded by an outstanding research team led by Nancy Nelson and Lina Shanley through the Center of Teaching and Learning at the University of Oregon. A special acknowledgement to my other team members who provided their expertise to help me succeed: Samuel Choo, Marcia Moore, and HyeonJin Yoon. Funding for the study was provided by the Institute of Education Sciences, Office of Special Education Programs (OSEP), through Grant H327S160019 to the University of Oregon. The opinions expressed are those of the authors and do not represent views of OSEP or the U.S. Department of Education.

A special shout-out to my colleagues from the school psychology and special education department at the University of Oregon for serving as math coaches and data collectors for the study. I am forever grateful to Steph Fisher, Mary Beth Hamilton, Tasia Brafford, Marah Sutherland, Britt Landis, David Fainstein, Abiola Taiwo, and Elizabeth Glenn.

My last acknowledgement is towards my dissertation chair and doctoral advisor, Dr. Hank Fien. I told him I wanted to complete a single case research design over evidenced-based educational technology, while evaluating both teacher and implementation practices and he believed in my abilities. I will be forever grateful for his investment in my professional growth.

To mothers who stay up late to help their children with their math homework and fathers who sign their papers when they failed to perform.

Robert Lloyd and Donna Marie, I owe you everything.

To my life partner, Austin. You believed in the dream before me and you sustained our life and our home throughout the whole process. I am looking forward to the next chapter of our life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table Page

I. INTRODUCTION

Math Achievement

According to NAEP, 40% of 4th-grade students are struggling to meet mathematical proficiency standards (NCES, 2016). Research supports that students who perform significantly low in mathematics will continue to struggle with basic math concepts throughout elementary school if early intervention is not applied (Morgan et al., 2009). Students who do not receive early intervention and continue to struggle with early mathematical skills throughout their development will have persistent challengers with mathematical concepts throughout their early adulthood (Geary, 2011). Early intervention and prevention methods are vital to mitigating the long term effects of persistent math difficulties.

Effective Teaching Practices in Mathematics

It is important for teachers to have a strong tier 1 math curriculum which targets the learning needs of students at risk for math learning difficulties. In addition to providing a strong tier 1 instructional program, schools need to provide supplemental instructional supports. Instructional programs in tier 2 and tier 3 need to provide intensive instruction targeting whole number concepts (Clarke et al., 2014). To understand how schools will effectively address the learning needs of students, they will need to have a comprehensive system of instructional support where intervention intensity can match the individual needs of the student (Clarke, et al., 2014). Teachers should have flexibility in their general classroom setting to intensify instructional supports based on continuous and on-going data collections efforts. To bridge the gap between general and special education services, schools need to be able to evaluate different service delivery models and the impact it has on student academic performance (Idol, 2006).

In addition to creating comprehensive service delivery models which address the needs of all students, teachers need to increase their use of explicit and systematic instruction during core math instruction (Doabler, et al., 2014). Explicit and systematic instruction during core instruction means teachers provide instructional models which ensure student accuracy, increased opportunities to practice critical math skills, and continuous feedback based off of student performance. In practice, teachers should use clear and consistent language when introducing new and complex mathematical content, provide academic corrective or affirmative feedback based on student responses, and increase the frequency in which they engage in mathematic verbalization during core math instruction to have the greatest impact on struggling learners in their classroom setting (Doabler, et al., 2014). When teachers engage in this process, they are increasing student practice opportunities for struggling student in their classroom setting. This process helps teachers become more familiar with their student's individual learning needs. One of the major goals of this current research study is to identify if there is a functional relationship between teacher use of technology-based interventions and the delivery of effective math practices during core math instruction like data-based decision-making and differentiated instruction.

Educational Technology

In 2013, Chueng conducted a meta-analysis of computer technology interventions and their impact on student math achievement in K-12 classroom settings. The effects of educational technology on student math achievement were enhanced for students with diverse learning needs, in elementary mathematics classrooms, and where a constructive approach to teaching was practiced (Chueng, 2013). This means educational technology programs can improve results of student math achievement in early elementary educational settings.

It is also important to note that intervention programs which were less than 6 months had a greater impact on student math achievement (Chueng, 2013). This means technology-based intervention programs which are supplemental to core math instruction may be better at improving student math achievement than longer and more intensive technology-based programs. When teachers decide to implement educational technology in their classroom, they should consider the degree to which experts have to review an instructional sequence and the current level of empirical evidence supporting the effectiveness of the intervention program (Nelson, Fien, Doabler, & Clarke, 2016). However, many technology-based intervention programs used in public school settings have not been independently evaluated by experts in the field (Nelson, et al, 2016). This makes it harder to determine which components of technology-based intervention programs are considered most useful by teachers and students. The number of features available through educational technology is high, but it’s important for research to narrow in on the factors which are important to the teacher and student use of technology-based programs.

There have been promising results about the use of educational technology to enhance student understanding of mathematical concepts. Educational technology provides teachers with the needed resources to customize student learning experience. Learner-centered approaches to teaching and student learning are gaining momentum in recent years. This is particularly true for educational technology intervention programs which incorporate direct instruction based on behavioral learning theory. When the instructional components of direct instruction are embedded in educational technology programs, it makes it easier for teachers and students to create individualized learning experiences (Kebritchi, 2008). To enhance game-based learning, developers need to report how educational technology is being used in their classroom settings and how the instructional materials are being integrated into typical teaching practices. It is important for the field to identify which instructional components of educational technology have the greatest effect on student engagement and teacher behavior (Kebritchi, 2008). Understanding the differences in implementation and use of educational technology will help explain why certain educational technology interventions are adopted by teachers and other programs are not.

Effective educational technology interventions include antecedents for desired behavior, active and frequent student responding, and feedback based on student responses (Moran & Malott, 2004). The way critical information on specific math skills is delivered to students is particularly important in technology-based intervention programs as well as in the general classroom setting. Educators need to consider how mathematical concepts are taught and introduced throughout an intervention program. Clear instructions and expectations make it easier for students to be successful when gaining new skills. It's also important to consider how students are prompted to engage with new academic content through the technology program. Prompting and fading of instructional demands is imperative because it helps the skills generalize across classroom activities (Moran & Malott, 2004). The last component to consider in educational technology is how the instruction can be adapted to serve a wide variety of classrooms and individual student's educational needs. Adaptive learning within educational technology makes it easier to present students with a new learning objective and provide scaffolded supports based on performance data (Moran & Malott, 2004). Adaptive and flexible learning makes it straightforward for the teacher to target individual skills within their classroom population.

Barriers to Implementing Educational Technology

Despite the promising results of educational technology, additional information is needed into how teachers select and implement technology-based programs in their classroom. To support the adoption of technology-based intervention programs in elementary school settings, it is important to consider the following pieces of information. First, intervention programs should help teachers identify the benefits and shortcomings associated with their technology program and how it is linked to curriculum and classroom instruction. Second, the design activities of the intervention program should focus on maintaining student engagement with meaningful character narratives. The last thing to consider is how the technology-based programs integrate both teacher and student involvement so they can each individualize their own learning experience (Okita & Jamalian, 2011). These three components on critical when developing technology-based intervention programs and should be critical design elements of the instructional program.

Changing Teacher Practices Through Educational Technology

A change in teaching practice through the use of a technology-based intervention could result in a change in student math achievement (Burns, et al, 2012). However, simply giving teachers access to instructional resources, although they are of high quality, will not increase teacher use of technology-based interventions (Grunwald Associates, 2010). If educational technology interventions are to be used effectively in school settings, researchers need to consider techniques and strategies which match the needs of both students and teachers. Both dynamics are considered when teachers adopt new teaching practices (Perlman and Redding, 2011).

Data-Based Decision-Making and Math Achievement

Data-based decision-making (DBDM) frameworks in educational settings often include four main components: analyzing student performance data, setting realistic goals for student growth, determining an instructional plan to address need, and executing the plan to evaluate effects (Van Geel et al., 2016). A DBDM process can be beneficial for students with learning disabilities because it supports early identification of student's needs and adoption of targeted instructional strategies (Stecker et al., 2005). DBDM can impact student math achievement, but it is often dependent upon the instructional setting (Van Geel et al., 2016). Interventions which incorporate DBDM have shown positive intervention effects for students from both low and high socio-economic status (Van Geel et al., 2016). However, DBDM systems are dependent upon teachers' ability to interrupt information accurately and teacher misinterpretations can often lead to less effective teaching practices and lower student math achievement growth over time (Van Geel et al., 2016).

Previous research studies have not been able to find a relationship between teacher use of data-based decision-making and differentiated instruction to support students with math difficulties. Additional information is needed to understand the relationship between these two educational practices and the impact it has on teacher and student outcomes (Faber, Glas, & Visscher, 2018). Observational data alone cannot accurately capture the process of data-based decision-making and differentiated instruction in authentic educational settings. It will be important to capture teacher's perspective and rationale about what student data is meaningful and what differentiated instructional practices are appropriate for supporting their individual students' math achievement (Faber et al., 2018). Large scale implementation of DBDM and differentiated instruction will require the use of instructional coaching to support teacher's skills like analyzing student performance data, setting achievable goals, and implementing appropriate instructional changes (Van Geel et al., 2016).

Development of NumberShire and NumberShire Integrated Tutor Systems (NS1-ITS)

The NS1 game was developed through the Institute of Education Sciences, Small Business Innovation Research Program as an educational technology game to help educators intervene early and strategically for students with or at risk for math learning difficulties using an explicit and systematic instructional sequence (Gause, Fien, Baker, & Clarke, 2011). Following the development of the NumberShire program, an implementation study was conducted to evaluate the feasibility and usability of the program in authentic educational settings. It was during this time that the NumberShire program received feedback from students and teachers to guide revisions. Once the revisions were incorporated into the intervention program, an 8 week abbreviated version of NumberShire was studied under rigorous experimental conditions to evaluate the promise it has for supporting student math achievement. Study results show the effectiveness of embedding educational technology with math interventions to improve early math outcomes for children with or at risk for math learning difficulties (Fien et al,. 2016). The study also indicated a need to ensure that these educational technology interventions are effectively designed and linked with teacher's day to day practices (Fien et al,. 2016).

These promising results led to the development of the NumberShire Integrated Tutor System (NS1-ITS). The NumberShire Integrated Tutor System is funded through an Office of Special Education research grant with the hopes of scaling up the NumberShire program to improve math outcomes for students with math learning difficulties. NS1-ITS includes an enhanced version of the NumberShire gaming intervention, data dashboard to support teacher use of data-based decision-making, instructional tutor, and implementation resource center (IRC).

The development of the NumberShire program and accompanying NS1-ITS teacher data dashboard is unique because it is based on the stage 3 implementation framework created by the National Institute of Health. This implementation framework targets both efficacy and effectiveness of intervention in real world settings under controlled experimental conditions (Onken et al., 2014). This research process allows for continued development of evidence-based programs, like NumberShire and NS1-ITS, while evaluating the utility and feasibility of use by teachers and students. These research studies are valuable because they help identify which components of the program are working as intended and which additional intervention components are needed to promote successful implementation. Implementation and treatment effectiveness must be evaluated simultaneously in order to transition intervention programs from controlled laboratory settings to real world application. When both factors are evaluated, it reduces the likelihood of intervention or implementation failure and increases the speed at which consumers are able to access evidence-based intervention programs (Proctor et al., 2011).

It is believed that when teachers engage with NumberShire Integrated Tutor System they will be able to engage in a sophisticated data-based decision-making process which will increase their use of differentiated and explicit instruction. This process will lead to an improvement in student learning behavior and math engagement in the classroom, thus resulting in improved math outcomes. Previous intervention studies which included a digital formative assessment and adaptive assignments tool have shown positive effects on student math achievement and motivation (Faber, Luyten, & Visscher, 2017).

Theory of Change

The Technology Acceptance Model (TAM) will be used to evaluate the perceived benefits of an enhance data dashboard and try to evaluate which technological components of the NS1-ITS program make it feasible for teachers to use educational technology to enhance their classroom instruction. TAM can help explain user motivation using three critical components: perceived usefulness, perceived ease of use, and attitude towards use. These factors determine how comfortable a teacher is with implementing new evidence-based technology into their typical teaching practices. The attitude of the user is explained primarily by perceived usefulness of a technological tool and ease of use, but also accounts for the behavioral intentions of the user (Taherdoos, 2008). The predictors used in the TAM are cognitive and focused on the adoption of actual behavior of a new technology to attitudes, beliefs, and perceptions of the intended user (Vanketesh & Bala, 2008).

Previous research studies which have used this model to identify math teacher's perception toward educational technology and the results indicated teacher's perception about the ease of use dominate their attitudes about technology-based programs (Nair & Mukunda, 2012). Math teachers also consider how useful the educational technology is for supporting their math instruction. One of the main findings from this study was that with adequate training and support, teachers were more likely to consider the tool useful and respond with a more positive attitude about implementing educational technology (Nair & Mukunda, 2012).

In addition to assessing the participants' internal perception towards the NS1-ITS program, we will also be capturing an aspect of their typical teaching behavior to identify if the teacher data dashboard has any impact on their daily teaching practices. It was important to capture teacher's perceptions, attitudes, and beliefs about educational technology, and how these different characteristics could impact the implementation of data-based decision-making and differentiated instruction. During the baseline stage of the intervention, we were able to capture teacher's intended use of educational technology in their classroom and compared the results to their teaching practices during the intervention phase.

II. METHODS

Research Question

The ultimate goal of this single-case research design is to determine a functional relationship between teacher's use of NS1-ITS and their use of data-based decision making (DBDM) and differentiated core math instruction for students with math difficulties in their general classroom setting. We would like to identify how first grade teachers would use the NS1-ITS teacher data dashboard and associated features to engage in data-based decision making and differentiated core math instruction. A secondary goal of this single case research design is to identify what NS1-ITS instructional resources 1st-grade teachers used most often to support student's understanding of whole number concepts. We will identify which instructional resources teachers used most often to individualize their students' instruction and discuss how they were utilized in their classroom settings. The proposed research question and hypothesis for the current study are listed below.

Research Question 1: What is the promise of NS1-ITS for improving teacher's use of DBDM when delivered in authentic education settings?

Research Question 2: What is the promise of NS1-ITS for increasing teacher's use of differentiated core math instruction designed to improve the quantity and quality of deliberate practice opportunities for students with or at risk for MLD delivered in authentic education settings?

It is hypothesized that using NS1-ITS will improve teacher's use of DBDM and differentiated core math instruction when delivered in authentic education settings as determined by classroom observations, teacher weekly self-reports, and frequency in which teacher access data dashboard features. It is also hypothesized that access to the NS1-ITS teacher dashboard will increase the opportunities teachers have to individualize their core math instruction based on student performance. It is believed that access to the enhanced NS1-ITS features will increase the quantity and quality of deliberate practice opportunities for students with MLD in authentic educational settings. The null hypothesis for the proposed study is that the NS1-ITS enhanced features will not indicate a change in teacher behavior when transitioning from the baseline to the intervention phase.

Research Question 3 for the present study will focus on the social validly of the NS1-ITS teacher data dashboard. It is it important to identify any possible barriers to implementing the NS1-ITS program in authentic educational settings. The team would like to evaluate if NS1-ITS is an appropriate intervention for supporting students with math difficulties. They will be focused on ease of use and perceived fit. Ease of use is defined as the participants ability to navigate the teacher data dashboard and deliver the NS1-ITS program as prescribed. Teacher's perceived use of the intervention is categorized by the perceived benefit the intervention would have for enhancing the math skills of students with math learning difficulties.

Research Design

The present research study will be using a multiple baseline across participants research design. This design was selected because the approach is versatile, easy to understand, and practical for work in school settings. Multiple baseline across participants designs are also useful when you have more than one subject seeking clinical treatment or when it is not possible to withdraw or remove the independent variable from the research design (Richards, 1999). Our goal is to demonstrate a functional relationship between teachers' use of the NumberShire Integrated Tutor System (NS1-ITS) and (a) teachers' use of data-based decision-making and (b) deliberate practice opportunities for students with math learning difficulties in authentic educational settings.

Research Timeline

This research study was conducted as a part of a bigger research project. The data collection for the current study took place from February of 2019 to June of 2019 in 2 public elementary school settings. Teacher and student pretesting took place in late February and continuous classroom observation data was collected from March of 2019 to May 2019. Post testing for teachers and student was conducted the first week of June.

All four teacher participants would begin using the NumberShire Level 1 game and accompanying teacher data dashboard for the first two weeks of the study. Once the two weeks were completed, teachers were randomly assigned to transition to NS1-ITS program. Teacher A started using the NS1-ITS program 2 weeks after the start of continuous data collection. Teacher B transitioned to NS1-ITS after 4 weeks of continuous data collection. Teacher C transitioned 6 weeks after continuous data collection began, and Teacher D would transition 8 weeks after continuous data collection had begun. Continuous data collection took place for 12 weeks in every classroom regardless of when the teacher transition to the NS1-ITS game and teacher data dashboard. Continuous data collection occurred Monday – Friday of each week. There was one week in April where continuous data were not collected because both school sites were closed for spring break.

When teachers transitioned to the NS1-ITS program they were provided with technical assistance in the form of a math coach. All of the math coaches were graduate students in the field of special education and school psychology. As a part of their training, math coaches attended three different training sessions where they were taught the different components of both NS1 and NS1-ITS. They received training on how to use the NS1-ITS instructional log. The goal of the math coach was to help teachers become familiar with the new enhanced features associated with the NS1-ITS program. Teachers would meet with their math coach once a week for 3 weeks to review student data and create instructional plans based on student performance through the NumberShire game. The math coach meetings would last between 30 to 45 minutes. The math coaches would use the NS1-ITS instructional log during their coaching sessions. The NS1-ITS instructional log included 4 main components associated with data based-based decision-making: (1) review available sources of student data, (2) identify trends, (3) locate available instructional resources for individualization, and (4) create an instructional plan. After the three weeks of instructional coaching were complete, teachers were able to use the NS1-ITS platform to review student performance data and create their individualized instructional plans independently. Appendix E details which instructional resources of the NS1-ITS program were developed prior to the start of the single case research time.

Teachers Participants

Four 1st grade teachers from two neighboring school districts located in the pacific northwest were recruited to participate in the research study. All of the teachers in the current research study were female. Out of the four teachers, two of the teachers had previous experience using NumberShire Level 1 (NS1) in their classroom setting. None of the teachers recruited to participate in the current study were familiar with the new enhanced features of the NumberShire Level 1 – Integrated Tutor System (NS1-ITS) program. The teacher participants all held four-year bachelor's degrees in the education field, while three of the four teachers had additional graduate training in education. The teachers had a large range of teaching experience which varied from 1 to 20 plus years of experience in the classroom. Active consent was acquired from all teaching participants and they were given a stipend of 250 dollars for their participation in the research study. Table 1 and 2 provides additional information regarding teacher and student demographics.

To participate in the study, teachers needed to have access to tablets and headphones for all participating students. The teachers agreed to administer the NumberShire program four times per week with each session lasting between 15 and 20 minutes. The teacher agreed to administer NumberShire regularly for over 12 weeks. In addition to implementing the NumberShire program, teachers had to identify five students in their classroom who needed additional support with whole number foundation skills. The student participants needed to be of the lowest-performing students on average (e.g., students who score below 20th percentiles in a classroom of 25 students). Teacher participants agreed to monitor their five selected students' progress on the NumberShire Level 1 and NS1-ITS game throughout the study. Parent consent and student assent were obtained for all the participating students to administer additional student educational assessments and conduct classroom observations of student-teacher interactions during core math instruction.

Setting – Core Math Instruction

It was important for researchers to capture what a typical 1st-grade classroom looked like during core math instruction to evaluate how educational technology could be used to enhance classroom instruction and influence teacher behavior. The participating teachers agreed to have classroom observations conducted during core math instruction at least 3 times per week for 12 weeks. Core math instruction was defined as the specific 1-hour block during the day in which math was the main instructional component being taught. Calendar time, which is typically considered math instruction, was excluded from the observation setting to focus on whole number foundation concepts addressed in the NumberShire program.

Classroom observations would take place for 15 minutes during the one-hour math block. Scheduled observations which exceeded 10 minutes of core math instruction were included in the study, scheduled observations which did not exceed 10 minutes of core math instruction were excluded and additional observations were scheduled. During the school week, the data collection team would observe different time frames throughout the same 1-hour block to capture every aspect of core math instruction. The varied observation start times helped create a comprehensive picture of core math instruction in each of the participating 1st-grade classrooms.

Independent Variable

NumberShire Level 1 (NS1) - Baseline

All participants will have access to the NumberShire Level 1 (NS1) game during the baseline stage of the research study. The original version of the NS1 educational game includes an accompanying teacher data dashboard. The NS1 student gameplay and associated teacher dashboard will both be used as a part of the baseline comparison. Including both components of the NS1 program, lets us evaluate how different parts of the intervention can be used to support student learning and teacher use of data-based decision-making simultaneously.

The NS1 educational game is set in an idyllic, Renaissance-themed world and features unique characters, storylines, and visual and reward content. Students have the opportunity to customize their game avatar and home base with fun costumes and household decorations. The items are awarded regularly when the student completes weekly sessions. NS1 consists of 12 hours of individualized instructional gameplay focused on building students' understanding of whole number concepts and skills identified in the CCSS-M, including place value concepts, fluency with number combinations, and word problem-solving. To build proficiency, students would spend approximately 15 minutes a day, 4 times per week, for 12 weeks to complete a total of 48 lessons.

NS1 provides supplemental math supports targeting whole number foundation skills. It is focused on increasing student math knowledge on CCSS-M in the following math domains: Counting and Cardinality, Operations and Algebraic Thinking, and Numbers and Operations in Base Ten. NS1 follows an explicit instruction model by providing students with scripted lessons and guided practice opportunities intended to ensure mastery of critical math content. The NS1 lessons include a warm-up activity, instructional model, assessment of previously taught skills, and wrap-up activity intended to reinforce critical math concepts for struggling learners (Fien et al, 2016). NS1 has been previously evaluated within the context of a randomized control study evaluating the intervention effects of NS1 for students with or at risk for math learning difficulties. Results determined students who participated in an NS1 treatment group outperformed students in the control condition on proximal measures of whole number concepts and skills (Fien, et al., 2016).

The NS1 teacher data dashboard helps teachers monitor student progress on the NumberShire Level 1 game. On the teacher data dashboard, student names are listed across the y-axis of your computer screen. On the x-axis, teachers can see all of the CCSS taught in the NS1 game. As students' progress through the NS1 game, the teacher data dashboard will start to fill with student performance data. Student performance data is immediately stored and updated on the teacher data dashboard once a student has completed their current lesson. If the student comes into contact with a previously taught skill, the dashboard will automatically update with the students' new scores. The student performance data presented on the dashboard is broken down into three categories which are associated with specific colors to broadly indicate how students are performing. Students who score above a 90% on a CCSSS will be displayed in green. Students who score between 89% and 75% on a CCSS will be displayed in yellow, while students who score below 74% on a CCSS are displayed in red. These three colors make it easy for teachers to evaluate student progress on the NS1 game and help teachers quickly identify gaps in student learning.

The NS1 teacher data dashboard displays NS gameplay data in multiple formats including a summary page, Common Core Report, student-level data, and item-level responses. This means teachers can view student performance data broadly from a classroom level perspective (using the summary page or Common Core Report) or take a more in-depth look at individual student performance and item-level responses. The summary-level data provides a broad overview of how everyone in the class is doing throughout the NS1 game. The summary report provides feedback about what math skills multiple students are mastering and which skills may require additional instructional support at the classroom level. The common core report lets teachers look at how students are progressing on CCSS-M specifically addressed in the NS1 game. This includes both kindergarten and 1st grade standards addressing Counting and Cardinality, Operations and Algebraic Thinking, and Numbers and Operations in Base Ten. The next way data is displayed on the teacher data dashboard is through the individual student report. This information allows teachers to review data for one of their students who is participating in the NS1 game. The data available at the student level provides insight into how one student is progressing across each standard they have completed. The individual student reports go more in-depth about the individual student strengths and weaknesses throughout the NS1 game and provide item-level responses about how a specific student is performing with a variety of whole number foundation skills. Teachers who use both classroom level data and individual student performance data can create a comprehensive evaluation of their students' current math ability within the context of the NS1 game.

The final piece of information available to teachers during the baseline stage is a weekly teacher report. Every week teachers would receive an email from a member of the NumberShire research team which would indicate their student's current lesson within the NS1 game. It would also provide information to teachers about where their students should be if they have been meeting their required dosage of 4 lessons per week.

This weekly teacher report helped teachers monitor their students' progression through the NS1 game.

NumberShire Level 1-Integrated Tutor System (NS1-ITS) - Intervention

The NumberShire Level 1 Integrated Tutor System (NS1-ITS) was developed to help teachers engage in data-based decision making and provide teachers supplemental instructional resources to support students' math achievement. The NS1-ITS platform includes an enhanced version of an evidence-based math game, called NumberShire Level 1 (NS1), and an accompanying teacher data dashboard that teachers can monitor student progress on the 1st grade Common Core State Standards in mathematics. The developers designed supplemental instructional resources to help teacher individualize their students' learning in their classroom setting through practice activities and instructional routines. Appendix C and D provides examples of the difference between NumberShire and NS1-ITS.

On the data dashboard, teachers can view and analyze student performance on the NS1 game, gain access to the Implementation Resource Center, (IRC) and technical supports. The IRC provides teachers and schools access to high-quality educational materials with strategies that can be incorporated into their daily teaching practices. The technological features of the NS1-ITS platform provide teachers resources to individualize their student gaming experiencing by repeating lessons or skipping mastered content. The aim of the teacher data dashboard is to make it easier for teachers to engage in data-based decision making and differentiate core math instruction by displaying student data, providing instructional recommendations based on student performance, and providing technological features which allow for individualized learning.

During the intervention phase of the study, teachers will have access to the NumberShire Level 1-Integrated Tutor System (NS1-ITS) game and accompanying teacher data dashboard. The NS1-ITS gameplay is similar to the NS1 game but includes enhanced features intended to promote student engagement. One of the major differences between NS1 and NS1-ITS is how the teacher and student can customize their gaming experience. Students are no longer required to move through the NS lessons in a linear trajectory. Teachers and students can individualize their own gaming experience by repeating previously taught lessons and skipping mastered skills. These enhanced gameplay features make it easier to provide flexible learning opportunities based off students' individual learning needs. In addition to providing flexible student gameplay options, students will now have accuracy incentives for correctly responding to test items. The student's NumberShire character will begin to glow different colors (silver to gold to a rainbow) when the student responds correctly to multiple test items during their math lesson. This accuracy incentives provide immediate feedback to the student while they are playing the NumberShire game. Other than these two main intervention components, NS1 and NS1-ITS student gameplay features are the same.

The major changes between NS and NS1-ITS occur on the teacher data dashboard. The NS1-ITS teacher data dashboard provides information about student progress on specific skills addressed in the NS1-ITS game. Once again student names are listed on the y-axis and CCSS math skills addressed in NumberShire are listed across the x-axis. Student gameplay performance data is immediately uploaded to the dashboard once a student has completed their daily lesson. In addition to generating student performance data, the NS1-ITS platform provides access to a wide variety of educational resources in the new Implementation Resource Center (IRC). The IRC is a tab located on the NS1-ITS teacher data dashboard and provides teachers with quick access to instructional resources. The IRC has four different categories which provide teachers with information on high-quality math resources and tips for implementing the NumberShire in their classroom. Additional information regarding their descriptions is listed below.

- *Best Practices in Math Instruction*: This section of the IRC provides additional information on best practices in math instruction. This section includes evidence-based teaching strategies, standards for math practice, and instructional recommendations for supporting diverse learners.
- *Multi-Tiered Systems of Support*: This section provides teachers with resources on how they can effectively implement MTSS for mathematics. Teachers will find information related to goal setting and examples of how to use data to make instructional decisions.
- *NS1-ITS Implementation:* In this section of the IRC, teachers can learn more information about the NumberShire program, including how to get started, the scope and sequence of the intervention program, and tips for implementing the program successfully.
- *Differentiated Instruction:* This is the last section listed in the IRC and its purpose is to help teachers individualize their classroom instruction based on student performance data. This section will provide teachers with

diverse instructional materials to address the needs of students in their classroom. It includes additional math practice activates and instructional routines to facilitate explicit teaching of whole number concepts.

Dependent Variables RQ 1

Weekly Teacher Survey and NumberShire User Log Data

To assess the promise NS1-ITS has for improving teacher's use of data-based decision-making in authentic education settings, we will be using two different types of data. The first piece of data used to evaluate this research question is the weekly Data-Based Decision-Making and Differentiated Core Math Instruction (DBDM and DCMI) weekly teacher survey. The DBDM and DCMI weekly teacher survey are based off a task analysis of the steps a teacher would need to take to evaluate their students' progress on the NumberShire game and make instructional decisions based on student performance. The DBDM and DCMI survey follow the same steps in the NS1-ITS instructional log. The instructional log is used by math coaches during their coaching sessions with teachers to evaluate student progress and identify additional instructional supports. A copy of the DBDM and DCMI survey is available in Appendix D.

Every Thursday, teachers would be emailed a Qualtrics link in which they were asked to answer questions based on the previous week's teaching behavior. The surveys would need to be completed by the following Monday. In total each teacher (n = 4) would complete 12 surveys based on the number of weeks they were implementing NumberShire. The weekly teacher survey is a self-report measure trying to capture the frequency in which teachers are assessing the teacher data dashboards and making instructional decisions based on student performance data. The DBDM and DCMI weekly teacher survey can be used to quantify the frequency and steps teachers followed regarding data-based decision-making and differentiated core math instruction. The range of scores on the DBDM and DCMI weekly survey teacher vary from 0 to 20 points depending upon teacher-reported responses. Higher scores on the DBDM and DCMI weekly survey indicate a higher level of teacher engagement in data-based decision making and differentiated core math instruction.

In addition to calculating the weekly teacher survey scores, NumberShire user log data will be used to describe the frequency in which teachers are logging on to the teacher data dashboard throughout the study. The NumberShire user log data will display the number of times per week a teacher logged into their NS1-ITS dashboard account. We will evaluate the DCMI and DBDM weekly teacher survey scores and number of teacher logins to discusses differences in how teachers were engaging in the data-based decision-making process to support students with math difficulties in their classroom.

Dependent Variables RQ 2

Classroom Observations Measures

To assess how educational technology can impact a teacher's delivery of differentiated core math instruction, continuous data was collected during core math instruction to track changes in teacher-student interactions. Two classroom observation tools were used to capture the quantity and quality of differentiated core math instruction. The first classroom observation tool used is the Classroom Observations of Student-Teacher Interactions (COSTI; Smolkowski & Gunn, 2012). The COSTI measures the rate of student-teacher interactions during core instruction, with particular attention to the number of teacher models provided, academic feedback of critical math content, and student practice opportunities. The operational definitions for these terms are listed below. The data collection team would observe for 15 minutes during core math instruction and conduct a frequency count. At the end of the 15-minute observation, they would calculate the total number of teacher models, academic feedback, and student practice opportunities.

- Teacher Models: Discrete teaching behavior. They are explicit, overt, and unambiguous explanation, communication, or physical demonstration of mathematical concepts. Teacher models can include definitions, step by step procedures, strategies or explanations, facts, or think aloud. A teacher models occur before a student response and must involve teaching. Questions are not considered teacher models.
- Academic Feedback: A teacher's verbal, physical, or written response to a group or individual student response. Academic feedback is overt, specific, and delivered immediately following a student response. There are two types of academic feedback including error correction and response affirming.
- Student Practices: Practice opportunities include verbal responses, use of math manipulatives, or written responses. Student practice opportunities can engage a group of students or individual students.

The second classroom observation tool used was the Ratings of Classroom Management and Instructional Support (RCMIS; Doabler & Nelson-Walker, 2009). The COSTI and RCMIS instruments have been field-tested and validated in more than 2000 classroom observations of reading and math instruction in elementary and middle school settings. The RCMIS is comprised of 4 items that target features of math instruction. This includes the instructional format, mathematics domains addressed, quality of instruction, and level of student engagement. Observers used a 6-point scale to rate the overall quality of each item (Doabler et al., 2012; Nelson-Walker, 2010). Effective classroom management, use of instructional time, active instructor monitoring, explicit and systematic instruction, and targeting critical content where all considered when evaluating the overall quality of instruction during classroom observations. Student engagement refers to the student attentiveness to the instructor and lesson and their participation in the assigned task and completion of those tasks. Quality of instruction was defined as "Quality of Instruction" refers to the use of evidence-based practices during lessons (e.g., effective classroom management strategies, focused use of instructional time, active instructor monitoring, explicit and systematic instruction targeting critical content). It was rated on a scale of 1 (very low quality) to 6 (very high quality). Student engagement refers to student attentiveness to the instructor and the lesson, participation in assigned tasks, and completion of activities. It was rated on a scale of 1 to 6 with 1 being no students engaged during any part of the lesson and 6 being all students engaged for the whole observation.

NumberShire Observation Form

Students who were selected by their teachers to participate in the intervention were observed once a week during their scheduled NumberShire time to see how the intervention program was delivered in authentic educational settings. There were two main purposes for collecting implementation fidelity data. The first was to compare the information to teacher self-report measures and tech data. The second was to control how students learned and what type of instruction they were receiving by either being in the baseline or intervention condition. Interrater reliability was gathered for 20% of the study and raters scored above 85% reliability across multiple observation sessions. NumberShire observations were conducted weekly and missed observation would be rescheduled for the same week.

Data collectors would gather the following pieces of information: observation date, start time, end time, school ID, Teacher ID, number of students, observer name, observation type (independent/practice), reliability-co observers. Scores ranged 0-24 with a 0-4 rating scale. There were 6 implementation questions rated using a scale of 1 (not applied) to 4 (consistently applied). Student Engagement was defined as student attentiveness to the instructor and the lesson, participation in assigned tasks, and completion of activities. Scale of 1 to 6 1 being no students engaged, 6 being all students engaged for all parts of the lesson.

Social Validity

To assess how comfortable teachers were with using educational technology and the NumberShire program, teachers were given two different surveys during pre and post-testing to identify how comfortable they felt using NumberShire in their classroom setting and how useful they perceived the intervention program to be at addressing their student's educational needs. The first survey conducted was a stage of concern survey (Appendix A). This survey was used to assess teacher's perception of both NumberShire programs, NS1 and NS1-ITS, and their perceived utility in their classroom setting. The survey helps identify teachers' perceptions of facilitators and barriers to implementing the intervention program as intended. The survey questions address the amount of time the participants can devote to learning a new intervention program and if they believe the NumberShire program will be able to accurately capture their students' math abilities. The survey includes 8 questions which participants will respond to on 4-point scale ranging from strongly agree to strongly disagree. A total score will be calculated during pre-and post-testing and the scores will be compared to see if there was a change in teacher perception on barriers and facilitators from baseline to intervention.

The final measure used during teacher pre and post-testing was the Intervention Rating Profile – 15 (IRP-15; Martens, Witt, Elliott, & Darveaux, 1985).The IRP-15 is a survey consisting of 15 items with responses on a 6-point Likert-type scale (1 = strongly disagree, 6 = strongly agree), with possible scores ranging from 15 to 90. Higher scores indicate a higher level of acceptability (social validity) towards NS1-ITS. Teachers complete their surveys at the beginning and end of the research study and their total scores will be compared to see if the difference in score is not due to chance alone. This measure will be used to identify teachers' perceptions of the current and enhanced teacher dashboards and the utility it has for supporting students with math difficulties. A copy of the survey is available on Appendix B.

RQ 1 and 2 Data Analysis

A standard mean difference approach will be used to determine the effect size of the NS1-ITS program on teacher's use of data-based decision-making and differentiated core math instruction. What Works Clearinghouse experts on single case design and analysis encourage the use of effect size calculators compared to visual analysis to evaluate intervention effects (Kratochwill et al., 2010). As such, the Tau-U non-parametric technique created by Vannest, Parker, Gonen, and Adiquzel (2016) will be used to evaluate the impacts of NS1-ITS on teacher behavior. Effect size scores .20 or less are considered to have a small effect, .20 to .60 have a moderate effect, .60 to .80 is considered to be a large effect, while a score of .80 or higher is considered to be a very large effect size.

Visual analysis will be discussed in addition to omnibus and individual effect sizes to discuss the impact NS1-ITS has on teacher models, academic feedback, student practice opportunities, and teacher self-reported use of data-based decision making and differentiated core math instruction. Descriptive data will be used to describe the quality and quantity of data-based decision making and differentiated core math instruction in authentic education settings. Descriptive data is important to single case design research studies because it provides an individualized perspective of how each teacher participant engaged in the data-based decision-making process and used the enhanced features of NS1-ITS to support students with math learning difficulties in their classroom.

The last pieces of information which will be discussed in the current research study is related to technology and implementation of NumberShire. The technology pieces of data will report on the number of logins and downloaded resource teacher used. Implementation data will discuss how the intervention was implemented each week in each of the teacher participants' classrooms. It will briefly describe students' level of engagement as well as the areas of implementation in which teachers were successful and the areas in which they required additional support. This information will beneficial when identifying a teacher's individual capacity to use educational technology with fidelity in their classroom setting.

RQ 3 Data Analysis

A reliable change index (RCI) will be used to evaluate the effects of the social validity measures provided by the teachers. It will evaluate whether a change over time of an individual score is statistically significant. The numerator represents an actual observed difference between scores and denominator represents a standard error of measurement. Ideally, the change in score will indicate statistical significance greater than a difference which have occurred due to random measurement error alone (Jacobson & Traux, 1991). A score of 1.96 or higher is considered to be statistically significant. Results will be shared for three out of four participants due to missing data points on posttest assessment measures.

III. RESULTS

Tau U Effect Size

138 classroom observations were conducted over the course of 12 weeks to evaluate the impact of NumberShire Level 1 Integrated Tutor System on teaching practices like Teacher Models, Academic Feedback, and Student Practices. An omnibus Tau-U score across participants for Teacher Models was 0.32 CI_{95} [0.11, 0.54]. This represents a moderate effect across all participants. The omnibus effect size for Student Practices was 0.21 CI_{95} [-0.0043, 0.43]. The results indicating a moderate effect size. There was no significant effect size reported for Academic Feedback or Weekly Teacher Survey. Additional information regarding omnibus effect size scores are available for reference on Table 3.

The Weekly Teacher Survey documenting teacher reported use of data based decision making and differentiated core math instruction reported large, positive effects for 2 out of the 4 teacher participants. Teachers B had a very large, positive and individual reported effect on their weekly teacher survey. Teacher D had a large positive effect size for their self-reported use of data-based decision making and differentiated core math instruction. Teacher B's calculated effect size was 1.00 CI_{90} [0.395, 1], while Teacher D had a calculated effect size of 0.71 CI_{90} [0.11, 1].

Visual Analysis of Classroom Observations and Weekly Teacher Surveys

Figure 1, 2, 3, and 4 displays the multiple baseline graph across participants for teacher models, academic feedback, student practices, and weekly teacher survey. All multiple baseline graphs are available to view at the end of this document in the appendix. When using visual analysis to review the impact of NS1-ITS on Teacher

Models, it is important to note that there is no clear trend in the data and the level does not dramatically change from baseline to intervention across the different teachers. The observation data is highly variable and there are multiple data points which overlap across the different phases. The intervention does not appear to have an immediate effect on the occurrence of teacher models during core math instruction. There appears to be a basic effect for immediacy and trend from baseline to intervention for Teacher B, but overall it did not demonstrate experimental control across all teacher participants.

The results are similar to the information regarding Teacher Models. There appears to be some basic effects regarding immediacy for Teacher C. However, it is difficult to identify a relationship regarding level and trend. The observational data appears consistent from baseline to intervention across participants and once again there is a consistent overlap of data points. There is not a demonstrated experimental effect on Academic Feedback across all teacher participants.

Student Practices demonstrates a different story regarding its visual analysis. There is a basic effect identified with Teacher D from baseline to intervention. The number of student practices greatly increases once teachers have access to the enhanced NS1-ITS features and there is a clear difference in level and trend. The number of student practices opportunities appears to be consistently above the baseline averages. There are only two overlapping data points between baseline and intervention. The remaining participants all have inconsistent data with no clear trend, level, or immediate effect. Almost all of the data points overlap for Teachers A, B, and C which means they did result in a significant change from baseline to intervention.

The most positive effects regarding visual analysis are demonstrated in the teacher's self-reported use of data-based decision making and differentiated core math instruction. Teacher A and C do not demonstrate a basic effect because of their inconsistency and trend level. Teacher C described less reported incidences of data-based decision making and differentiated core math instruction during intervention. Teacher C is demonstrating a negative trend on their self-reported measure and the data points overlap with baseline data. Teacher A appears to be moving in a positive trend towards week 12 of the study, however additional data points would be needed to confirm this hypothesis. Teachers B and D demonstrate a positive level and trend. There is a basic effect demonstrated for Teacher B and D. It is easy to see a difference in teacher reported scores from baseline to intervention on the Weekly Teacher Survey measure. It appears that Teacher B had a greater effect on self-reported incidences of data-based decision making and differentiated core math instruction than Teacher D, but they both report a clear change in teacher practices from baseline to intervention.

Classroom Observations and Weekly Teacher Survey

Table 4 displays the baseline and intervention scores for all teacher participants regarding classroom observations and their weekly teacher survey. Overall, Teacher Models increased for 3 out of the 4 teacher participants. Teacher A's number of teacher models stayed the same from baseline to intervention, while Teachers B, C, and D increased during the intervention phase. There were no significant changes regarding Academic Feedback across all teacher participants. During intervention, the teachers demonstrated less than or equal to their baseline scores on Academic Feedback. For Student Practices, Teacher D increased their score from 30.6 during baseline to 67.69

during intervention. Teachers C and B remained the same, while Teacher A had a decrease in student practices during the intervention phase.

In addition to providing the baseline and intervention scores, Table 5 indicates the average number of Teacher Models, Academic Feedback, and Student Practices per minute by each teacher participant. On average, participants provided 1 Teacher Models per minute during core math instruction across baseline and intervention. They averaged less than one Academic Feedback per minute during a 15-minute classroom observation. Teacher participants provided 1 Student Practice per minute during baseline and averaged 2.38 Student Practices during intervention.

Additional information was gathered regarding student engagement and quality of instruction. Student engagement levels remained the same for Teacher A whose baseline average was 3.13 and intervention average of 3.17. Teachers B, C, and D all demonstrated a small positive change regarding student engagement. Once again, Teacher A had a decrease in their quality of instruction while Teachers B, C, and D all made positive improvements in this area. Table 6 indicates the baseline and intervention scores for student engagement and quality of instruction.

Descriptive Data for DBDM

Teacher A

Teacher A data sources reported using student game play data and iReady Quiz Reports to evaluate her students' math achievement and progress towards CCSS. Throughout the course of the study, Teacher A decreased in their reported ability to facilitate NumberShire, have students complete a NumberShire lesson, access teacher data dashboard, and reviewing of individual student data. They believed that their available student data was not representative of their students' math ability. They reported three trends over the course of 12 weeks on their weekly teacher survey. All the reported trends identified by teacher participants are available to reference in Appendix F. Overall, Teacher A reported concerns regarding her student's ability to understand that each successive number is 1 more.

Teacher A would frequently review student mastery data on the teacher data dashboard. In the beginning of intervention, they access the following components on the implementation resource center (IRC): differentiated instruction best practices in math instruction, and NS1-ITS implementation. They did not review the Multi-tiered Systems of Support section until the very end of the study. This section includes resources about how to build your MTSS process within your school system. Teacher A would frequently access the student account information to have students repeat a NumberShire lesson. They also reviewed the NS1-ITS storyline and alignment with the CCSS. Throughout the study, Teacher A accessed 2 different instructional routines targeting add 1 and grouping ones. They also reviewed 7 different student practice worksheets which aligned with the NumberShire student game play.

Teacher B

Teacher B would often use informal assessment like EasyCBM and student game play data to monitor their students' math achievement. Teacher B reported good facilitation of NumberShire during baseline, however, it decreased during intervention. Teacher B had an increase in their students' ability to complete a NumberShire lesson, access the teacher data dashboard, and review individual student data. Teacher B had no change from baseline to intervention regarding the accuracy of their students' math scores. They identified 2 trends during baseline and 7 trends during the intervention phase.

Once Teacher B gained access to the NS1-ITS teacher data dashboard, they immediately reviewed the Differentiated Instruction section of the IRC. Throughout the study, they accessed over 8 different instructional routines targeting counting, comparing groups and/or numbers, and fluently adding and/or subtracting. They also accessed over 15 different student practice activities and would use them across multiple weeks to support student learning. At the end of the study, they reviewed all of the available resources including student account information which lets them individualize student gaming experience. Overall, Teacher B was much more likely to differentiated their math instruction using instructional routines and practice activities sheets. They were less likely to change their students' gaming experience and would provide them additional opportunities to practice important skills.

Teacher C

On their Weekly Teacher Survey, Teacher C only reported viewing student performance data for 2 out of the 12 weeks of the study. When they were reviewing their students' math achievement, they would only use student game play data. They reported only one trend throughout the course of the study. They had an increase in their facilitation of NS and completing of lessons, but they decreased in their report of accessing the teacher data dashboard and review of individuals student data. During the intervention phase of the study, they believed that the available math data represented their students' true math ability despite the limited review of available student information. Their baseline average on this area was 3.67 while their intervention average was 4.75.

Teacher C stopped accessing the teacher data dashboard after the 10th week of the study. They first reviewed student mastery data and account information when they gained accessed to the teacher data dashboard before reviewing the Differentiated Instruction materials of the IRC. In terms of NS1-ITS materials, they accessed 6 student practice activities and 4 instructional routines. The instructional routines selected included support with counting by 1, adding fluently, comparing numbers, and grouping tens. It appears that it was difficult for Teacher C to maintain a DBDM and DCMI framework throughout the duration of the study. They seemed to have relied heavily on the support of their instructional coach in order to access the NS1-ITS features and implement the instructional resources regularly in their classroom. Despite access to high quality educational materials, Teacher C did not show much change in their ability to individualize instructional supports for students with math difficulties.

Teacher D

Teacher D was similar to Teacher B and used informal assessment like easyCBM and student game play data. They increased on the following measures from baseline to intervention: facilitation of NumberShire, completing a NumberShire lesson, accessing of the teacher data dashboard, and review of individual student data. Regarding the accuracy of available student math data, Teacher D reported an average of 3.8 during baseline and an average of 4.75 during intervention. C and D rarely identified trends on their weekly teacher survey, however, teacher D would frequently review student performance data while Teacher C did not.

Teacher D accessed 13 different student practice activities. Out of those 13 practice activities, she repeatedly accessed 5 of them over the course of 3 weeks. They also used 5 different instructional routines which targeted counting objects, fluently adding and subtracting, comparing numbers, equal signs with numbers and objects, and subtracting 10. From the Implementation Resource Center, Teacher D would review student mastery data and the differentiated instruction section of the teacher data dashboard.

NS1-ITS Tech Data

Information was gathered over the number of times, teacher participants would login to the teacher data dashboard. Overall, teachers were able to log in at least once per week to evaluate student progress on the NS1-ITS game. Teacher A and C had an average of 1 log in per week during intervention. Teacher B had an average of 2.37 logins per week, while Teacher D had an average of 1.5 logins per week.

The research team was also curious about how teachers navigated the NS1-ITS teacher data dashboard and accompanying resources. Information was gathered over the links teachers would click on when navigating the teacher data dashboard. Teacher D had the highest number of clicks during intervention with an average of 11.25 clicks per week. Teacher A was not far behind with an average of 10.3 clicks per week. Teacher B had 8.5 clicks per week and Teacher C had an average of 5.82 clicks per week during intervention.

Result findings indicated mixed levels of engagement regarding data analysis and accuracy of student performance. More specifically, teachers B and D consistently reviewed all student performance data throughout the study and believed it accurately represented their students' mathematical understanding; and teacher A and C infrequently accessed the teacher data dashboard and there was a decrease in their review of individual student performance data throughout the course of the study. On the Weekly Teacher Survey, Teachers B and D had higher levels of self-reported DBDM and DCMI than Teachers A and C whose scores actually decreased throughout the duration of the study.

All participating teachers in the study were able to individualize their instruction and design individual learning objectives for students using recommendations from the NS1-ITS. Teachers B and D frequently accessed and delivered the NS1-ITS instructional routines and student practice worksheets to support individual student learning objectives. These resources are embedded into the teacher data dashboard and can be easily linked with student performance data. Teachers A and C repeated NumberShire lessons, reviewed math vocabulary, or used concrete manipulatives to individualize their instruction. Other than repeated NS lessons, these individualized learning opportunities often required additional time and resources for the teacher to implement successfully.

Implementation of NS1-ITS

The teacher participants all had various degrees of implementation when they delivered NumberShire. Teacher A was consistently good at getting their students started with the NumberShire program. Teacher B had an increase in their overall score from baseline to intervention suggesting that it was easier for them to implement the program when using the NS1-ITS features and teacher data dashboard. In addition to starting the session appropriately, Teacher B was effective at managing her students' equipment (headphones, tablets, usernames and passwords). Teacher C and D were also effective at managing their student's equipment.

There were some challenges observed during the administration of NumberShire. Teacher A show a reduction in scores from baseline to intervention on the following domains: appropriate equipment, monitoring student during the game, assuring completion of lessons, and having an established routine when students finish early. Their overall implementation score dropped from an average of 23 during baseline to 20.13 during intervention. There was also a reduction in their student engagement levels. During baseline student engagement had an average rating of 5 and during intervention it was 3.75.

Teacher B had an improvement regarding their ability to implement the NumberShire program. Their baseline average was 23 and increased to 24.29 during intervention. Teacher B struggled with monitoring students while they were playing NumberShire, establish a routine for students who finished early, and assuring completion of NumberShire lessons. They were able to make improvements during intervention regarding an effective start procedure and using appropriate equipment. Unfortunately, their student engagement levels decreased throughout the course of the study. During baseline, student engagement was rated as 5.25 and it dropped to 4.2 during intervention.

Teacher C struggled with monitoring students while they participated in NumberShire, implementing an effective procedure for when students finish early, and assuring the completion of NumberShire lessons. They did increase their overall implementation of the educational technology program. Their baseline average was 20 and intervention average was 23.8. Their student engagement remained relatively the same from baseline (4.5) to intervention (5).

Despite a small decrease in implementation from baseline (27.28) to intervention (26.2), Teacher D was highly successful at implementing the NumberShire program with fidelity. The only area in which Teacher D saw a reduction in scores was having an effective start procedure. Teacher D was the only participant to utilize the support of educational assistant during the implementation of NumberShire. The educational assistant support may have resulted in higher scores regarding implementation because students were consistently monitored, tech issues were addressed immediately, students had an established routine for when they ended early, and the educational assistant could assure completion of NumberShire lesson. During this time, Teacher D was able to support other students in her classroom while simultaneously individualized instructional supports for students with math learning difficulties.

Research Question 3 – Social Validity

A reliable change index (RCI) was used evaluate the difference in teacher scores on pre and posttest measure and to determine if the change in scores is statistically significant or due to random chance. Reliable change indexes use psychometrics components of their outcome measures in order to evaluate change over time for the same participant. The numerator represents an actual observed difference between scores and the denominator represents the standard error of measurement. Ideally, the change in score will indicate statistical significance greater than a difference which have occurred due to random measurement error alone (Jacobson & Traux, 1991). Scores which are above a 1.96 threshold are considered to be significant. Teacher B's reliable change index could be reported due to missing data on their post-test assessments.

Stages of Concern Survey

A small sample size of five was used to calculate the Cronbach alpha reliability score and standard error of measurement for the stages of concern survey. The survey had a standard deviation of 3.81 with a Cronbach's alpha of .63. The standard error of measurement was calculated at 2.33. Teacher A had a 2 point difference in score between pre and post testing. They had a reliable change index of 0.86 which is not consider significant. However, Teacher C had a score difference of 11 points between pre and post testing. They received a reliable change index score of 4.73. This reliable change index is above the 1.96 threshold and is considered significant. Teacher D did not demonstrate a change in scores between pre and post testing and therefore it was not possible to calculate a reliable change index. Teacher C was the only participants in the study to have a significant change in their ratings from baseline to intervention on the stages of concern survey.

At the end of the study, teachers were a lot more aware of the associated features of NS1-ITS and knew how to implement NS1-ITS in their classroom. Overall, teachers reported a positive change in their ability to learn and implement NS1-ITS. However, there was an overall reduction in their score regarding their perceived ability to gather additional resources or support for implementing the NS1-ITS program. This information suggest that any supplemental instructional resources meant to support student math achievement should be quickly and easily assessible for teachers using educational technology. If supplemental instructional supports are embedded into educational technology, teachers will spend less time reviewing and selecting supplemental resources and more time supporting targeted student learning objectives.

One of the next stages of development for NS1-ITS is an instructional tutor which will provide recommendations for individualized instruction based on student performance data. The information gathered from the Stages of Concern survey indicate a need for supplemental resources which can be quickly be linked to student performance data for individualized instructional support.

Intervention Rating Profile-15

The present study used the Cronbach's alpha of .98 and teacher standard deviation of 19.8 as defined by Marten et al., (2004). These psychometric components were used to calculate the reliable change index for all participants who had complete data sets all pre and posttest measures. Teacher A had a pre-post test score difference of 11 and yield a reliable index of 2.78. Teacher C had a difference of 18 between pre and post testing and obtained a reliable score index of 4.55. Teacher A and C yield significant results indicating that the change in score was not due to random chance alone. Teacher D yield a 7-point difference in pre and post test scores on the Intervention Rating Profile 15 and obtained a reliable change index of 1.77, which is below the statistical significance threshold and is not consider a reliable change in scores from pre to post testing.

It is important to note that Teacher D rated NumberShire positively during their baseline assessment which could explain why there was not a significant change in their rating of the NumberShire Level 1 Integrated Tutor System from baseline to intervention on the Intervention Rating Profile-15. Teachers A and C were less familiar with the NumberShire program at the start of the study and had a greater potential for growth than Teachers B and D because they both had previous experience with the NumberShire program. The information suggest that Teachers A and C considered NumberShire to be a socially significant intervention for supporting students with math learning difficulties in their classroom. Their scores on the assessment during baseline were lower than scores during intervention suggesting that the features associated with NS1-ITS were rated more favorably.

Tables 7 and 8 provided information regarding the reliable change index scores for the Stages of Concern Survey and Intervention Rating Profile. Table 9 and 10 detail the overall percentages scores for the social validity measures. Teachers reported needing additional support regarding their abilities to implement NumberShire Level 1 Integrated Tutor System in their classrooms. Teacher scores on the stages of concern survey range from 64% to 81% during baseline and 72% to 87% during intervention. Teachers reported NumberShire as a valid measure on the Intervention Rating Profile with an acceptability level ranging from 68% - 84% in baseline and 81% to 100% when provided with the enhanced technology features. Despite some doubts in their ability to implement NumberShire in their classrooms, teachers reported NumberShire as a favorable intervention for supporting students with math learning difficulties.

IV. DISCUSSION

The quality and quantity of data-based decision making and differentiated core math instruction were evaluated using classroom observations, teacher weekly reports, and the frequency of teacher logins and downloaded resources on NS1-ITS data dashboard. When considering the classroom observations, there was a change in teacher behavior from baseline to intervention for three out of four participants on teacher models. Teachers averaged 1 teacher model per minute during both baseline and intervention. Academic feedback did not have a change in score from baseline to intervention for any participant. Teachers averaged less than 1 incident of academic feedback per minute during baseline and intervention. This data suggest that it may be more difficult for teachers to provide immediate academic feedback during core math instruction. Only one teacher participant had a significant change for student practice opportunities. This creates a stronger argument for educational technology programs which use an explicit instruction model to help ensure student success when learning new instructional material. Students may not always receive academic feedback during core math instruction and educational technology could supplement those opportunities within the general education setting.

On the weekly teacher survey, teachers reported their frequency of data-based decision making and differentiated instruction for student with math difficulties. There where was a reduction in the quality and quantity of data-based decision making and differentiated core math instruction for two out of the four participants. One of the teachers struggled to effectively implement NumberShire with their students and it may have impacted their engagement with the enhanced features which support data-based decision making and differentiated core math instruction. The last participant had a significant effect on their self-reported use of data-based decision making and differentiated core math instruction. There appear to be a distinct difference in scores for Teachers A and C and Teacher B and D. Teachers B and D have each had previous experience with implementing NumberShire in their classroom. The non-positive effect scores for Teacher A and C may be due to their individual capacity to deliver a new educational intervention program while simultaneously evaluating student performance data and creating individualized instructional plans.

All of the teachers were able to log into the teacher data dashboards at least once per week throughout the 12-week study. Teachers B and D utilized the data dashboard more frequently than Teachers A and C. Once again, this difference in teacher behavior could be due to previous exposure to NumberShire. Teachers A and C received minimal training on how to implement NumberShire in their classroom and would have benefitted from more exposure to the scope and sequence of the intervention prior to the implementation of data-based decision making and differentiated instruction.

Teachers B and D demonstrated higher scores on their implementation of NumberShire suggesting they benefitted from additional time and exposure to educational technology before implementing teaching practice which support student learning. Teachers A and C received the same level of instructional coaching as Teachers B and D. However, it seems like they were unable to maintain the prescribed level of data-based decision making and individualized instruction required by the research team. It is possible that there would be different results in their quality and quantity of data-based decision making and differentiated instruction if the number of students there were monitoring were smaller or the frequency of data-based decision making and differentiated instruction was greater than once a week.

It would be important for researchers and student support teams to identify how teachers want to implement educational technology in their classroom especially since teachers reported concerns with their own ability to independently implement NumberShire outside of other teaching requirements. Teacher D who has the most successful implementation across all participants, utilized the support of an educational assistant to be able to ensure completion of all NumberShire lessons. Teacher' perspective and expertise should always be considered when thinking about their rate of data-based decision making and individualized instruction for students of various educational needs. It is possible for general education teachers to adopt new teaching practices like data-based decision making and individualized instruction especially important when working in inclusive classroom settings.

Overall, the results from the current study show that it is possible for educational technology interventions like NumberShire to influence teaching behavior during core math instruction. Educational technology which incorporate student performance data, engage students in the learning process, and consider both teacher and student factors should be priority when adopting new evidence-based practices (Okita & Jamalian, 2011). Additional information is still needed about the appropriate frequency of data-based decision making and individualized instruction a teacher would need to engage in to make improvements on student math achievement. The current rate of once per week seems to be a demanding timeline for teachers to review student performance data, prescribe an instructional plan, implement supports, and evaluate any change in performance.

Implementation

When researchers use implementation frameworks, they address both efficacy and effectiveness of educational technology interventions (Onken et al., 2014); thus, reducing the likelihood of intervention or implementation failure and increasing the speed at which consumers are able to access evidence-based programs (Proctor et al., 2011). Researchers should consider the instructional skills required to successfully implement DBDM and evaluate supplemental supports, like the ones described in this study.

A single case research design provides researchers and practitioners with a broader perspective on the various ways' teachers adjusted their instruction to support their students' educational needs through the use of evidence-based, educational technology. Single case research designs make it possible to analyze individual differences in order for successful integration of educational technology in classroom settings, researchers and school officials must consider what implementation drivers contribute to the success of intervention programs in their intended setting (Sims and Melcher, 2017).

The current study highlights the importance of training and coaching to support teachers who are in the process of developing effective and enhanced teaching practices, like data-based decision making and differentiated instruction, through the use of educational technology. Teachers A and C were able to successful report student performance data and create instruction plans when they were provided with an instructional coach to support their skill development. Once the coaching was removed, it was difficult for both teachers to maintain the same level of data-based decision making and differentiated core math instruction. Teachers who begin using a new program will require more help to effectively navigate educational technology. Building an effective response to intervention framework using educational technology will require administrative support which provides ongoing technical assistance to teachers throughout the adoption of new evidence-based practices. This step will be highly important to address in order to ensure successful implementation, integration, and adoption of educational technology in classroom settings.

Limitations

When using a single case research design, it is impossible to account for any external stimuli which could impact the quality of results. Time of school year, end of the year assessments requirements, transitions to summer vacation, repetition of previously taught mathematical concepts, and teacher motivation could all have impacted the intervention effects for the current research study. It would be meaningful for future researchers to think about implementing this study at the beginning of the school year when these extraneous variables are not as powerful and teachers' motivation to provide supplemental intervention support to students with math difficulties may be higher. It would be interesting to see if time of year has an impact on teacher use of educational technology and their willingness to engage in data-based decision making and differentiated instruction. Understanding the difference between implementing educational technology at the beginning or middle of the school year, would be an important factor for school teams to consider when selecting and adopting educational technology which fits within their response to intervention frameworks.

Another limitation to the current study is the interrater reliability of 85% across all classroom observations. Despite the fact that this level of reliability is considered to be appropriate for single case research designs, the visual analysis results made it difficult to detect any intervention effects. Intervention results were able to be evaluated through an effect size analysis, but visual analysis alone would not have been enough to detect a significant change from baseline to intervention. Future researchers may want to increase their interrater reliability to 90% or higher in order to evaluate the intervention effects of educational technology on effective teaching practices like teacher models, academic feedback, or student practice opportunities. A higher level of interrater reliability would reduce the likelihood of unrelated factors impacting intervention results and would help establish a relationship between educational technology interventions and effective teaching practices.

It's possible that a smaller intervention effect occurred because teacher models, academic feedback, and student practice opportunities are not explicitly taught through the teacher data dashboard and enhanced features of the NumberShire Level 1 Integrated Tutor System. Like many other educational technology tools which are available, NumberShire provides detailed information about student performance, but at this stage of development it did not prescribe teachers with specific rules on how to engage in data-based decision making and differentiated instruction. The weekly teacher survey had more positive results on teacher behaviors than the classroom observations. This may have occurred because the weekly teacher survey was a more proximal measure of effective teaching practices than the distal measure of classroom observations. It will be important for future researchers to consider both distal and proximal measures when evaluating the intervention effects of educational technology using a single case research design.

Lastly a larger sample is preferred for calculating the reliable change index score for the stages of concern survey. The population for the current study (n = 4) was used to identify the reliability and standard error of measurement needed, but caution is warranted when evaluating the change in scores from pre and post testing.

APPENDIX A

THE STAGES OF CONCERN SURVEY USED DURING TEACHER PRE AND POST TESTING TO IDENTIFY BARRIERS AND FACILITATORS TO IMPLEMENTING EDUCATIONAL TECHNOLOGY IN THEIR CLASSROOM SETTING

Please rate how much you agree or disagree with the statements below:

	Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree	Strongly Agree
I am interested in using the NumberShire Integrated Tutor System to support math learning.	○	○	○	○	○	○
I have students who will benefit from the NumberShire Integrated Tutor System.	○	○	○	○	○	○
I am aware of the features of the NumberShire Integrated Tutor System.	○	○	○	○	○	○
I know what I need to do to facilitate the use of the NumberShire Integrated Tutor System in my classroom.	○	○	○	○	○	○
I believe that the NumberShire Integrated Tutor System will help my students learn math.	○	○	○	○	○	○
I believe that the NumberShire Integrated Tutor System will generate information about my students' math learning that I can use in my class.	○	○	○	○	○	○
I have time to learn and implement the NumberShire Integrated Tutor System.	○	○	○	○	○	○
I have other resources to implement the NumberShire Integrated Tutor System	○	○	○	○	○	○

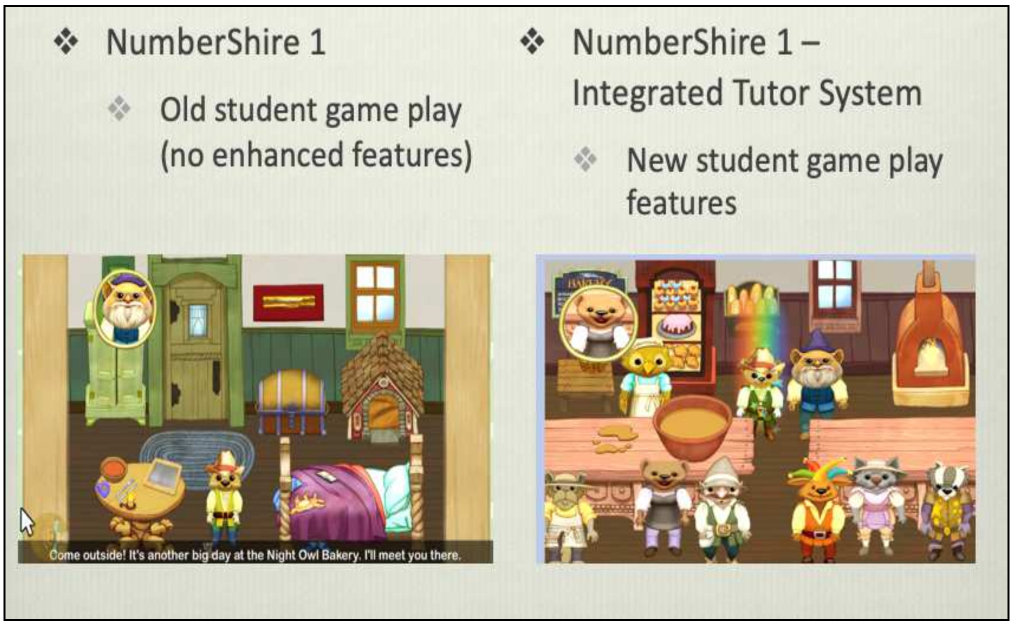
APPENDIX B

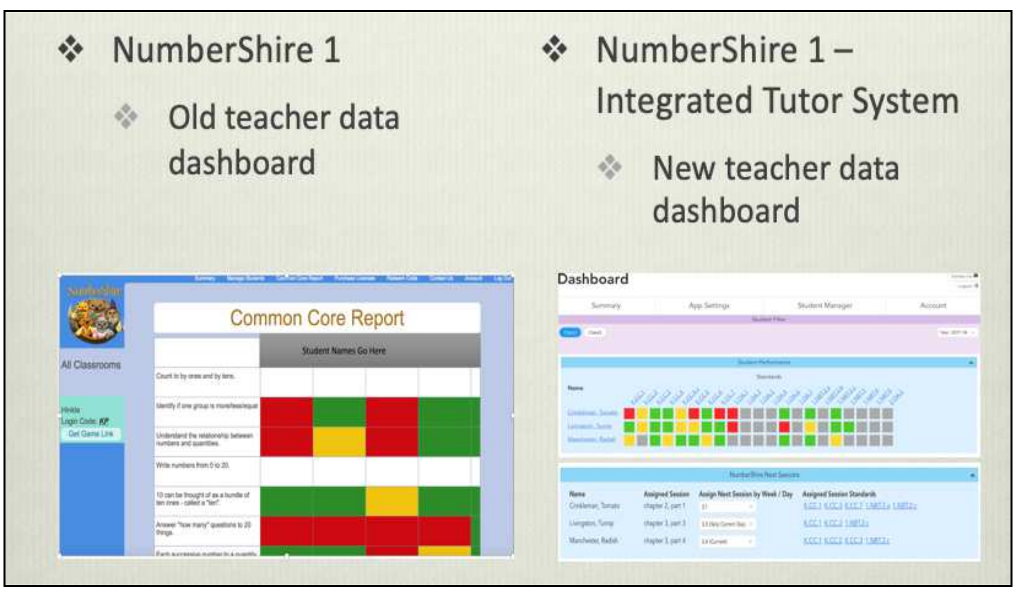
THE INTERVENTION RATING PROFILE-15 USED DURING TEACHER PRE AND POST TESTING AS A SOCIAL VALIDITY MEASURE FOR USING NUMBERSHIRE TO SUPPORT STUDENTS WITH OR AT RISK FOR MATH LEARNING DIFFICULTIES

	Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree	Strongly Agree
This intervention would be an acceptable intervention for my student's needs.	○	○	○	○	○	○
Most teachers find this intervention appropriately for students with similar needs.	○	○	○	○	○	○
This intervention should prove effective in supporting the student's needs.	○	○	○	○	○	○
I would suggest the use of this intervention to other teachers.	○	○	○	○	○	○
My students' needs are severe enough to warrant use of this intervention.	○	○	○	○	○	○
Most teachers would find this intervention suitable for the needs of their students.	○	○	○	○	○	○
I would be willing to use this intervention in the classroom setting.	○	○	○	○	○	○
This intervention would not result in negative side effects for my students.	○	○	○	○	○	○
The intervention would be appropriate for a variety of students.	○	○	○	○	○	○
The intervention is consistent with those I have used in classroom settings.	○	○	○	○	○	○
The intervention is a fair way to handle the student's needs.	○	○	○	○	○	○
The intervention is reasonable for the needs of the student.	○	○	○	○	○	○
I like the procedures used in this intervention.	○	○	○	○	○	○
This intervention would be a good way to handle my student's needs.	○	○	○	○	○	○
Overall, the intervention would be beneficial for my students.	○	○	○	○	○	○

APPENDIX C

FIRST IMAGE: SIDE BY SIDE COMPARISONS OF THE NS1 STUDENT GAME PLAY AND NS1-ITS STUDENT GAME PLAY. THE NS1-ITS PICTURE SHOWS ACCURACY INCENTIVE. SECOND IMAGE: SIDE BY SIDE COMPARISON OF NS1 TEACHER DATA DASHBOARD AND NS1-ITS TEACHER DATA DASHBOARD





APPENDIX D

DBDM AND DCMI WEEKLY TEACHER SURVEY

	100% (2 points)	1-99% (1 point)	0% (0 points)
1. I facilitated NumberShire in the classroom this week.			
1. My students completed a NumberShire lesson.			
2. I accessed the teacher report/teacher data dashboard.			
3. I looked at individual student's data.			

2. I used the following pieces of student/game play data (Check all that apply):

   Student Game Play Data on Weekly Teacher Report (1 point)
   Student Game Play Data on Teacher Dashboard (1 point)
   Other student assessment data (examples: informal assessment, easyCBM): (open response) (1 point for each piece of data listed)
   I did not use student/game play data this week. (0 points)
   Note any trends you saw: (open response)

3. On a scale on 1 (very poor) to 5 (great), I considered the data to be accurately capture my student's math ability.

4. I adjusted instruction using (teachers will check all the sources they used)

   a. Instructional Routine (1 point)

   b. NS1-ITS Practice Worksheet (1 point)

   c. Repeat NumberShire Lesson (1 point)

   d. Used concrete examples/manipulatives (1 point)

   e. Review math vocabulary (1 point)

f. Fluency probes (1 point)

g. Flash cards (1 point)

h. Gave them an additional assignment/homework (1 point)

    Specify:

i. Used another evidence-based curriculum: (1 point)

    i. Which curriculum?

j. Other: (fill in the blank) (1 point)

k. I did not adjust instruction (0 points)

Weekly Teacher Points:

________/________

APPENDIX E

NS1-ITS Implementation Resource Center (IRC) Overview for SCD Study

Category	Description	Completed
Best Practices in Math Instruction	Learn more about best practices in math instruction, including evidence-based teaching strategies, the Standards for Math Practice, and approaches for supporting diverse learners.	• Common Core State Standards in Math Standards • IES practices guides
Differentiated Instruction	Find tools to help you meet the varied instructional needs of your students, including additional math practice, instructional routines to facilitate explicit teaching of whole number concepts, and resources to support students' self-regulation skills.	• Instructional Routines • Practice Activities • Repeated NS Lessons
Multi-Tiered Systems of Support	Get resources to help you effectively implement MTSS for mathematics, including tools for setting goals and using data to inform instruction.	• MTSS Video
NS1-ITS Implementation	Learn more about NumberShire, including how to get started, the scope and sequence of the intervention, and implementation troubleshooting tips.	• NS1 Story Line • NS1 and CCSS alignment

APPENDIX F

Participant	**Week**	**Identified Trend**
Teacher A	2 – BL	Students are struggling with understanding that each successive number is one more.
Teacher A	3 - IV	Students are struggling with the concept of knowing the next number on the number line is one more...
Teacher A	5 - IV	I've seen a lot of improvement in counting on by 1 and knowing that each successive number is 1 more.
Teacher B	2 - BL	I was surprised by a lower than expected performance for one or two of my target students with skills that they show mastery in everyday tasks. I watched STUDENT working on NumberShire today and noticed that they seemed to be impulsive with answering. I made a point of sitting with them to encourage taking more time to think about the questions and being more careful with their answers. I made sure they noticed the feedback that the game was giving them with their performance.
Teacher B	3 – BL	Noticed one student was not accessing NumberShire lessons as often as the rest of their target group. I checked in with them and found that they were having some difficulty with independently logging in. I also noticed that the skills demonstrated during game play and EasyCBM math progress monitoring assessment are lower than classroom performance tasks. My opinion is that these students approach tasks impulsively and bit rushed without applying true abilities and skills. Learning to slow down and work carefully with more precision will be something that needs to be emphasized during instruction.
Teacher B	5 -IV	Three students are struggling with addition and subtraction with numbers 1-20
Teacher B	7 - IV	Fewer students in red/yellow but still struggling in same areas.
Teacher B	8 - IV	Same areas of strength and weakness in addition subtraction with number 0-20, as well comparing numbers.

Teacher B	9 - IV	-Counting objects from 1-20 continues to really stink with most students. -STUDENT is super high risk across the board and I think they are just clicking away quickly without really engaging. - Target kids are struggling with adding and subtracting with numbers 1-20 Adding and subtracting within 100 is red for all target students (no big surprise there).
Teacher B	10 - IV	Target kids are all struggling with the meaning of the equal sign as well as addition/subtraction with numbers 0-20
Teacher B	11 - IV	It seems that most kids have been more successful with picture representations of numbers. Most students did very well with identify groups of tens and ones given pictures. Exercises with 2-digit numbers and demonstrating the understanding of what the number in the tens or one's place means needs more practice.
Teacher B	12 – IV	A couple of students are plowing through lessons and with little success and I like 1-3 of them to go back and review and repeat those lessons.
Teacher C	3 - BL	That students struggle with identifying one more and one less.
Teacher C	5 - BL	Identifying one more and one less

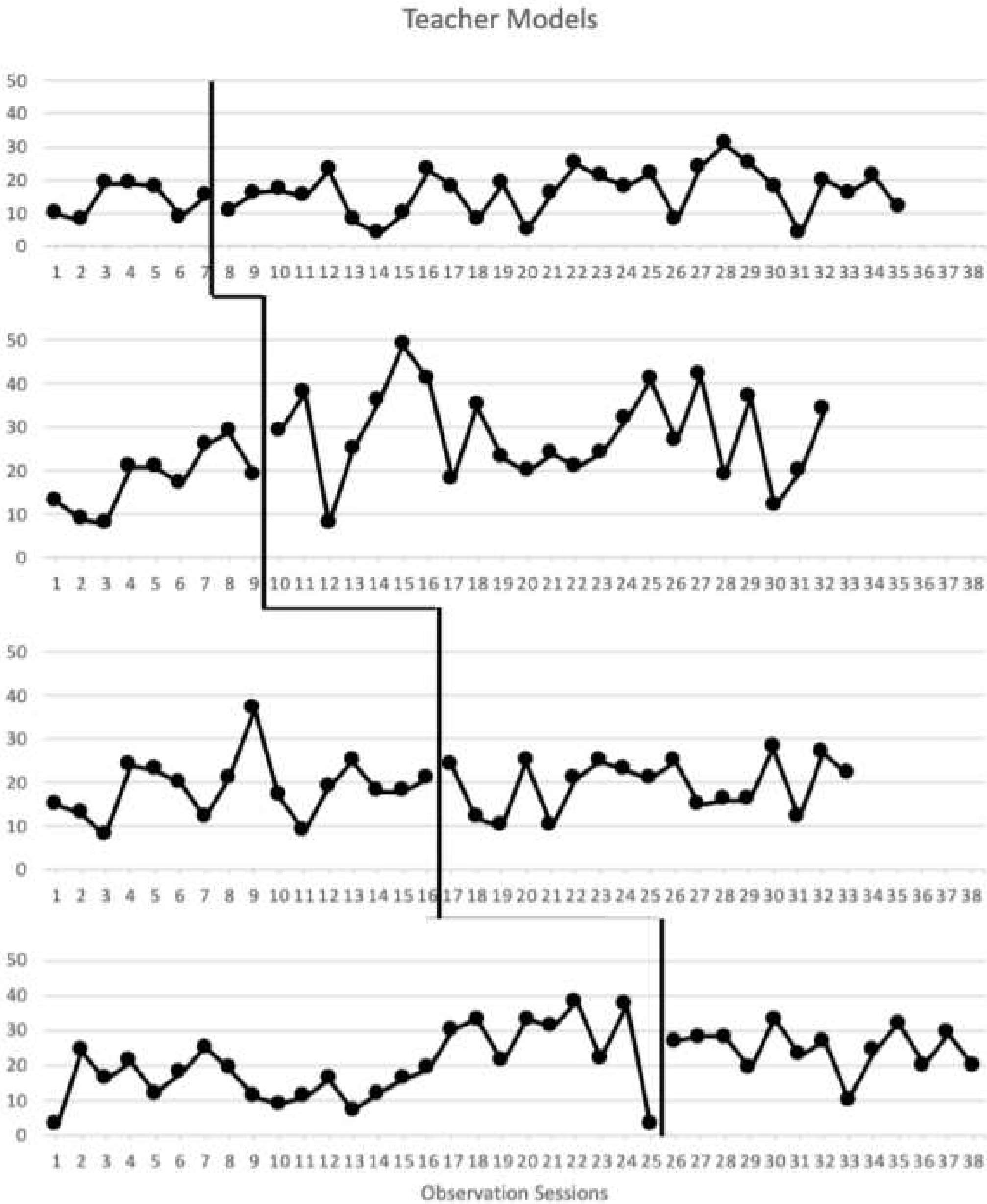


Figure 1. Multiple Baseline Across Participants Graph for Teacher Models by Class Observation

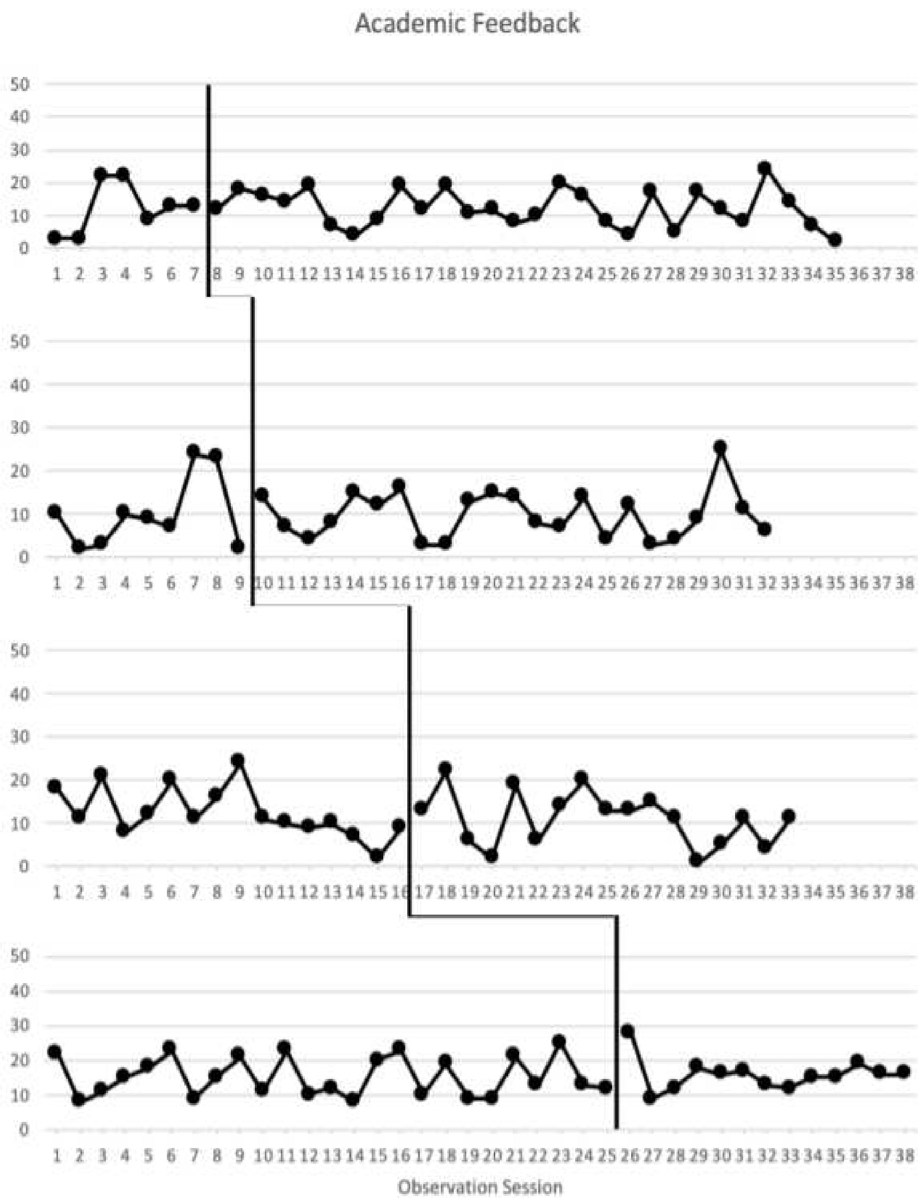


Figure 2. Multiple Baseline Across Participants Graph for Academic Feedback by Observation Session

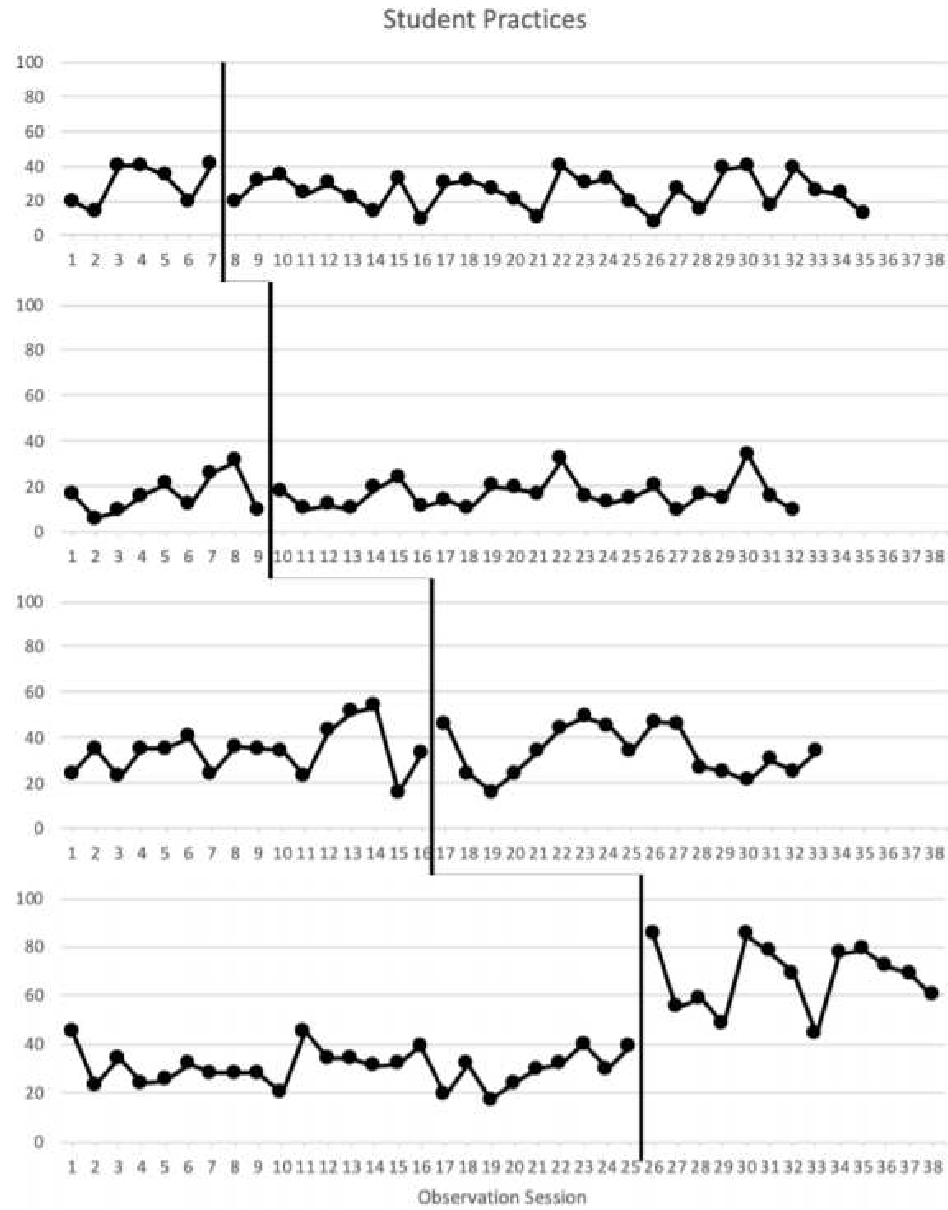


Figure 3. Multiple Baseline Across Participants Graph for Student Practice by Observation Session

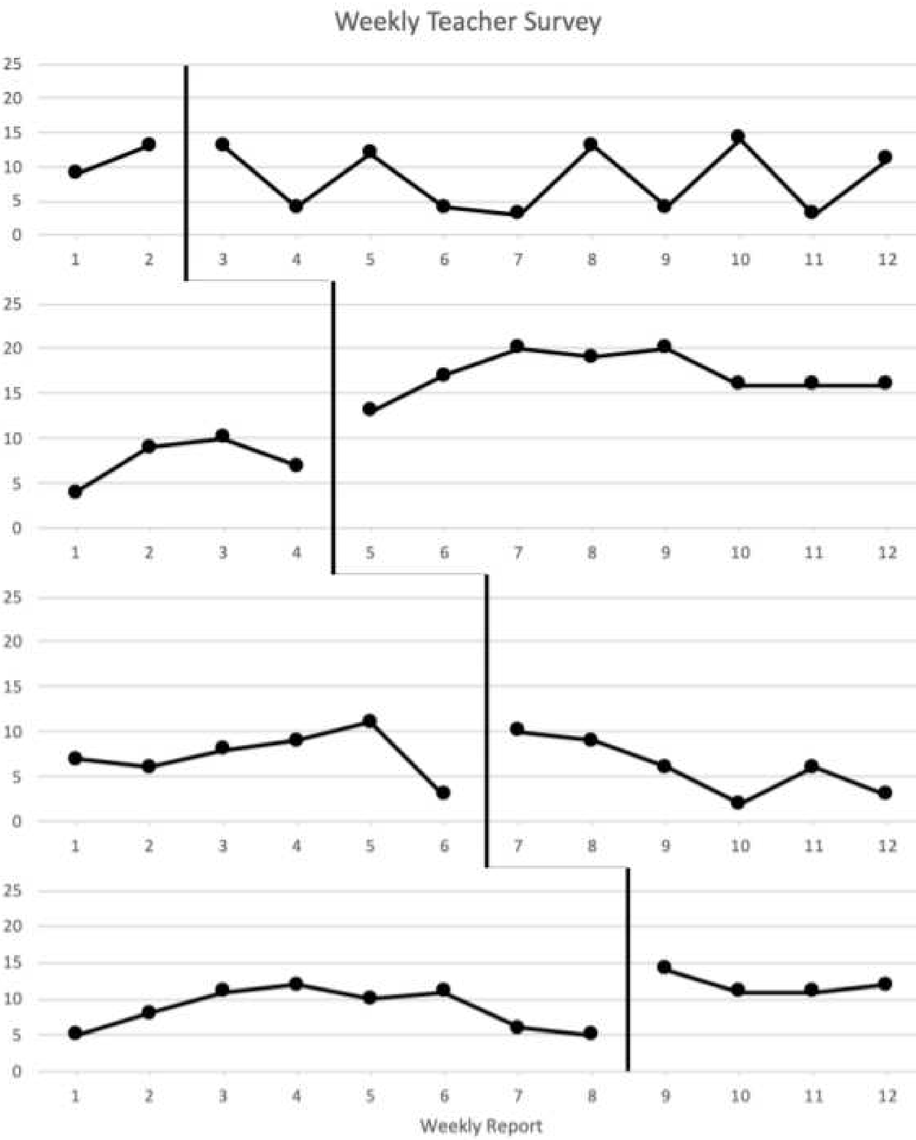


Figure 4. Multiple Baseline Across Participants Graph for Teacher Reported Use of Data Based Decision Making and Differentiated Core Math Instruction by Wee

Table 1

Teacher Demographics by Teaching Experience, Graduate Training, and Certifications

	Teacher A	*Teacher B*	*Teacher C*	*Teacher D*
Teaching Experience	1 – 5 years	20 + years	10 – 20 years	1 – 5 years
Previous Experience with NumberShire	No	Yes	No	Yes
Graduate Training	NA	Yes	Yes	Yes
Certifications	NA	SPED and ELL	NA	ESOL

Note. NA stands for not applicable, SPED stands for special education, ELL stands for English Language Learner, and ESOL stands for English as a Secondary Language.

Table 2

Student Demographics by Classroom Teacher

	Teacher A	*Teacher B*	*Teacher C*	*Teacher D*
Number of Students	5	5	5	5
School	A	B	A	B
Gender				
Male	1	2	1	2
Female	4	3	4	3
Ethnicity				
White	5	5	5	4
Hispanic/Latino	0	0	0	1

Note. Student participants were the lowest-performing students on average in their classroom (e.g., students who score below 20th percentile in a classroom of 25 students).

Table 3

Omnibus Effect Sizes of Classroom Observation and Weekly Teacher Reports during Core Math Instruction

	Tau	*p-Value*	*Z Scores*	*95% C.I*	*Degree of Effect*
Teacher Models	0.32	0.00**	2.96	0.11 - 0.54	Moderate
Academic Feedback	0.46	0.68	0.41	-0.17 - 0.26	Small
Student Practices	0.21	0.05*	1.92	-0.00 - 0.42	
Weekly Teacher Survey	0.25	0.26	1.26	-0.13 - 0.63	

Table 4

Baseline Averages and Intervention Averages for Classroom Observations and Weekly Teacher Reports of Data-Based Decision Making and Differentiated Core Math Instruction by Participant

	Teacher Models	*Academic Feedback*	*Student Practices*	*Weekly Teacher Survey*
Teacher A				
Baseline	16.33	14.17	34.83	11
Intervention	16.36	12.28	25.17	8.1
Teacher B				
Baseline	18.11	10	16.33	7.5
Intervention	28.47	9.8	16.52	17.13
Teacher C				
Baseline	18.75	12.43	33.81	6.83
Intervention	19.53	10.94	33.59	6.5
Teacher D				
Baseline	19.48	15.2	30.6	8.5
Intervention	24.61	15.85	67.69	12

Note. Teachers B and D have had 1 full year of NumberShire compared to Teachers A and C who do not have previous teaching experience with the intervention.

Table 5

Average Number of Teacher Models, Academic Feedback, and Student Practices Per Minute Across Baseline and Intervention by Teacher Participant

	Teacher Models	*Academic Feedback*	*Student Practices*
Teacher A			
Baseline	1.09	0.94	2.32
Intervention	1.09	0.81	1.67
Teacher B			
Baseline	1.21	0.66	1.08
Intervention	1.89	0.65	1.10
Teacher C			
Baseline	1.25	0.82	2.25
Intervention	1.30	0.72	2.24
Teacher D			
Baseline	1.29	1.01	2.04
Intervention	1.64	1.05	4.51

Note. Teachers B and D have had 1 full year of NumberShire compared to Teachers A and C who do not have previous teaching experience with the intervention.

Table 6

Baseline Averages and Intervention Averages for Student Engagement and Quality of Instruction by Teacher Participants

	Student Engagement	*Quality of Instruction*
Teacher A		
Baseline	4.14	3.13
Intervention	3.70	3.17
Teacher B		
Baseline	4.00	4.40
Intervention	5.10	4.81
Teacher C		
Baseline	4.31	3.80
Intervention	4.68	4.05
Teacher D		
Baseline	4.19	5.11
Intervention	5.11	5.14

Note. Scores are based on a 1 (very low) to 6 (very high) scale. Teachers B and D have had 1 full year of NumberShire compared to Teachers A and C who do not have previous teaching experience with the intervention.

Table 7

Difference in Scores and Reliable Change Index by Teacher Participant on Stages of Concern Survey

	Teacher A	*Teacher B*	*Teacher C*	*Teacher D*
Difference in Scores	2	-	11	0
Reliable Change Index	0.85	-	4.73*	0

Note. Scores above 1.96 are considered significant. Results for Teacher B cannot be reported due to missing data on post-test assessments.

Table 8

Difference in Scores and Reliable Change Index by Teacher Participant on Intervention Rating Profile 15 (IRP-15)

	Teacher A	*Teacher B*	*Teacher C*	*Teacher D*
Difference in Scores	11	-	18	7
Reliable Change Index	2.78*	-	4.55*	1.77

Note. Scores above 1.96 are considered significant. Results for Teacher B cannot be reported due to missing data on post-test assessments.

Table 9

Percentages in Total Scores by Teacher Participants on the Intervention Rating Profile 15 (IRP-15)

	Teacher A	*Teacher B*	*Teacher C*	*Teacher D*
Baseline Rating	68%	84%	80%	80%
Intervention Rating	81%	-	100%	87%

Note. Teacher B did not complete her posttest assessments and an intervention score on the Intervention Rating Profile 15 could not be reported. Teachers B and D had previous experience with the NumberShire program.

Table 10

Percentages in Total Scores by Teacher Participants on Stages of Concern Survey

	Teacher A	*Teacher B*	*Teacher C*	*Teacher D*
Baseline Rating	68%	70%	64%	81%
Intervention Rating	72%	-	87%	81%

Note. Teacher B did not complete her posttest assessments and an intervention score on the Stages of Concern Survey could not be reported. Teachers B and D had previous experience with the NumberShire program.

REFERENCES CITED

Burns, M. (2012). Effect of a Computer-Delivered Math Fact Intervention as a Supplemental Intervention for Math in Third and Fourth Grades. Remedial and Special Education., 33(3), 184-191.

Cheung, A. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. Educational Research Review., 9(1), 88-113.

Clarke, B., Doabler, C. T., Nelson, N. J., & Shanley, C. (2015). Effective instructional strategies for kindergarten and first-grade students at risk in mathematics. *Intervention in School and Clinic*, *50*(5), 257-265.

Doabler, C., & Nelson-Walker, N. J. (2009). Ratings of classroom management and instructional support. Unpublished observation instrument, Center on Teaching and Learning, College of Education, University of Oregon, Eugene.

Doabler, C. T., Nelson, N. J., Kosty, D. B., Fien, H., Baker, S. K., Smolkowski, K., & Clarke, B. (2014). Examining teachers' use of evidence-based practices during core mathematics instruction. *Assessment for Effective Intervention*, *39*(2), 99-111.

Geary, D. (2011). Cognitive Predictors of Achievement Growth in Mathematics: A 5-Year Longitudinal Study. *Developmental Psychology*, 47(6), 1539-1552.

Grunwald Associates. (2010). *Educators, technology and 21st century skills: Dispelling five myths. A study on the connection between K–12 technology use and 21st century skills.* Bethesda, MD: Author. Retrieved from http:// www.grunwald.com/pdfs/Educators_ Technology_21stCentury-Skills

Faber, J., Glas, C., & Visscher, A. (2018). Differentiated Instruction in a Data-Based Decision-Making Context. *School Effectiveness and School Improvement*, 29(1), 43-63.

Faber, J., Luyten, H., & Visscher, A. (2017). The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment. *Computers & Education*, 106, 83-96.

Fien, H., Doabler, C. T., Nelson, N. J., Kosty, D. B., Clarke, B., & Baker, S. K. (2016). An examination of the promise of the NumberShire Level 1 gaming intervention for improving student mathematics outcomes. *Journal of Research on Educational Effectiveness, 9*(4), 635-661. doi: 10.1080/19345747.2015.1119229

Idol, L. (2006). Toward inclusion of special education students in general education: A program evaluation of eight schools. *Remedial and Special Education*, *27*(2), 77-94.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19.

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*. Retrieved from What Works Clearinghouse website: https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_scd.pdf

Kebritchi, M. (2008). Examining the pedagogical foundations of modern educational computer games. *Computers & Education*, *51*(4), 1729-1743.

Martens, B. K., Witt, J. C., Elliott, S. N., & Darveaux, D. X. (1985). Teacher judgments concerning the acceptability of school-based interventions. *Professional psychology: Research and practice*, *16*(2), 191.

Morgan, P. L., Farkas, G., & Wu, Q. (2009). Five-Year Growth Trajectories of Kindergarten Children With Learning Difficulties in Mathematics. *Journal of Learning Disabilities, 42*(4), 306-321. doi:10.1177/0022219408331037

Moran, D. J., & Malott, R. W. (Eds.). (2004). *Evidence-based educational methods*. Elsevier.

Nair, I., & Mukunda Das, V. (2012). Using Technology Acceptance Model to assess teachers' attitude towards use of technology as teaching tool: A SEM Approach. *International Journal of Computer Applications, 42*(2), 1-6.

Nelson, N. J., Fien, H., Doabler, C. T., & Clarke, B. (2016). Considerations for realizing the promise of educational gaming technology. *Teaching Exceptional Children*, *48*(6), 293-300.

Okita, S., & Jamalian, A. (2011). Current Challenges in Integrating Educational Technology into Elementary and Middle School Mathematics Education. Journal of Mathematics Education at Teachers College, 2(2), 49-58.

Onken, L. S., Carroll, K. M., Shoham, V., Cuthbert, B. N., & Riddle, M. (2014). Reenvisioning clinical science: Unifying the discipline to improve the public health. *Clinical Psychological Science*, *2*(1), 22-34.

Perlman, C. L., & Redding, S. (2011). Handbook on Effective Implementation of School Improvement Grants. *Academic Development Institute*.

Richards, S. (1999). Single subject research : Applications in educational and clinical settings. San Diego: Singular Pub. Group.

Sims, B., & Melcher, B. (2017). Active implementation frameworks: Their importance to implementing and sustaining effective mental health programs in rural schools. In *Handbook of rural school mental health* (pp. 339-361). Springer, Cham.

Smolkowski, K., & Gunn, B. (2012). Reliability and validity of the Classroom Observations of Student–Teacher Interactions (COSTI) for kindergarten reading instruction. *Early Childhood Research Quarterly*, *27*(2), 316-328.

Stecker, P. M., Fuchs, L. S., & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of research. Psychology in the Schools, 42(8), 795–819.

Taherdoost, H. (2018). A review of technology acceptance and adoption models and theories. *Procedia Manufacturing*, *22*, 960-967.

Van Geel, M., Keuning, T., Visscher, A., & Fox, J. (2016). Assessing the Effects of a School-Wide Data-Based Decision-Making Intervention on Student Achievement Growth in Primary Schools. *American Educational Research Journal, 53*(2), 360-394.

Vannest, K.J., Parker, R.I., Gonen, O., & Adiguzel, T. (2016). Single Case Research: web based calculators for SCR analysis. (Version 2.0) [Web-based application]. College Station, TX: Texas A&M University. Available from singlecaseresearch.org

Venkatesh, V., & Bala, H. (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions. Decision Sciences, 39(2), 273-315.