Journal of School Psychology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of School Psychology

journal homepage: www.elsevier.com/locate/jschpsyc

Using assessment to individualize early mathematics instruction

Carol McDonald Connor[a,*], Michèle M.M. Mazzocco[b], Terri Kurz[c], Elizabeth C. Crowe[d], Elizabeth L. Tighe[e], Taffeta S. Wood[a], Frederick J. Morrison[f]

[a] University of California, Irvine, United States
[b] University of Minnesota, United States
[c] Arizona State University, United States
[d] University Academy, Panama City, FL, United States
[e] Georgia State University, United States
[f] University of Michigan, United States

ARTICLE INFO

Action Editor: Albers Craig

Keywords:
Mathematics
Elementary Education
Instruction
Individualize Child Differences

ABSTRACT

Accumulating evidence suggests that assessment-informed personalized instruction, tailored to students' individual skills and abilities, is more effective than more one-size-fits-all approaches. In this study, we evaluate the efficacy of Individualizing Student Instruction in Mathematics (ISI-Math) compared to Reading (ISI-Reading) where classrooms were randomly assigned to ISI-Math or ISI-Reading. The literature on child characteristics X instruction or skill X treatment interaction effects point to the complexities of tailoring instruction for individual students who present with constellations of skills. Second graders received mathematics instruction in small flexible learning groups based on their assessed learning needs. Results of the study (n = 32 teachers, 370 students) revealed significant treatment effects on standardized mathematics assessments. With effect sizes (d) of 0.41–0.60, we show that we can significantly improve 2nd graders' mathematics achievement, including for children living in poverty, by using assessment data to individualize the mathematics instruction they receive. The instructional regime, ISI-Math, was implemented by regular classroom teachers and it led to about a 4-month achievement advantage on standardized mathematics tests when compared to students in control classrooms. These results were realized within one school year. Moreover, treatment effects were the same regardless of school-level poverty and students' gender, initial mathematics or vocabulary scores.

1. Introduction

Children who develop strong mathematics skills in their early school years generally experience greater long term academic achievement than do their peers with poor numeracy skills (Duncan et al., 2007; Siegler et al., 2012), which in turn predicts later school completion and occupational success, (Geary, Hoard, Nugent, & Bailey, 2013; National Mathematics Advisory Panel, 2008; Ritchie & Bhatia, 1999; Rivera-Batiz, 1992; Siegler et al., 2012). However, according to the National Assessment of Educational Progress (NAEP, 2015), only about 40% of United States 4th graders attain proficiency in mathematics, and this percentage drops to 24% for children living in poverty. These rates represent an enduring problem, because achievement levels have held fairly steady over the last decade (NAEP 2005 to 2015). This means that a significant number of current and future U.S. students will enter school on a path towards underachievement in mathematics. At issue, then, is the need to identify ways to reverse this path and to do so as

* Corresponding author.
E-mail addresses: connorcm@uci.edu (C.M. Connor), mazzocco@umn.edu (M.M.M. Mazzocco), Terri.Kurz@asu.edu (T. Kurz), ecrowe@university-academy.org (E.C. Crowe), etighe@gsu.edu (E.L. Tighe), tswood@uci.edu (T.S. Wood), fjmorris@umich.edu (F.J. Morrison).

http://dx.doi.org/10.1016/j.jsp.2017.04.005
Received 11 May 2016; Received in revised form 12 April 2017; Accepted 20 April 2017
0022-4405/

early as possible because early mathematics underachievement does not simply correct itself with time (Aunola, Leskinen, Onatsu-Arvilommi, & Nurmi, 2002; Mazzocco & Myers, 2003). One way, we hypothesize, is to improve the early mathematics instruction students receive, so that it is better aligned with their individual learning strengths and weaknesses in general, and their mathematical knowledge and skills specifically. However, the features of such individualized learning needs and instruction (as they pertain to early mathematics) are not yet well understood, and identifying them has been proposed as one of the current "grand challenges" of mathematical cognition research (Alcock et al., 2016, p. 23). In this paper, we describe an approach to mathematics instruction that uses assessment results to individualize (i.e., personalize, differentiate) the instruction students receive. We then test the efficacy of this instructional approach in a randomized controlled trial with teachers randomly assigned to individualized student instruction in either mathematics or reading.

1.1. *Influences on early mathematics achievement*

Accumulating evidence suggests that differences in mathematics achievement levels are related to multiple sources of influence (Berch & Mazzocco, 2007; Bronfenbrenner & Morris, 2006) including family, community, and student characteristics. These sources of influence are also likely to influence mathematics development and children's kindergarten entry quantitative skills, which are highly predictive of later mathematics achievement (e.g., Nunes, Bryant, Evans, & Barros, 2015; Purpura, Baroody, & Lonigan, 2013). Although early mathematical skills include spatial and geometric skills, much of the research on early mathematical thinking focuses on number (e.g., Mazzocco & Räsänen, 2013), and number is a primary focus for early mathematics in the Common Core State Standards (CCSS, Common Core State Standards Initiative, 2010). There is evidence of individual differences emerging from studies of very basic, primitive measures of magnitude judgment (Halberda, Mazzocco, & Feigenson, 2008) to highly formalized representations of number relations and the ability to fluently execute simple operations (Nunes et al., 2015). For instance, kindergartner's understanding of enumeration, symbolic numbers, and the relations between numbers are strong predictors of their math achievement level at the end of third grade (Mazzocco & Thompson, 2005) or fourth grade (Jordan, Glutting, & Ramineni, 2010; Morgan, Farkas, & Wu, 2009). Their ability to compose and decompose numbers at school entry predicts numeracy levels in high school (Geary et al., 2013). Moreover, individual differences persist throughout the school age years, when they may manifest as qualitative errors (e.g., place value errors, atypical computational errors; Mazzocco, Murphy, Brown, Rinne, & Herold, 2013; Mazzocco, Myers, Lewis, Hanich, & Murphy, 2013) or differences in fluency versus accuracy (Petrill et al., 2012; Price, Mazzocco, & Ansari, 2013). Importantly, these differences are apparent in early childhood (e.g., Desoete, Ceulemans, De Weerdt, & Pieters, 2012; Murphy, Mazzocco, Hanich, & Early, 2007), which means that we cannot assume children arrive at school with equivalent foundational mathematics skills.

Even if children do enter school with solid foundational mathematics skills, the path to their mathematics competence may not be maintained throughout primary school. Moreover, efforts to promote mastery are likely to support achievement only for those children who have not yet achieved it, and efforts to maintain mastery will do little to advance mathematics achievement of students who have already attained age-appropriate early mathematics skills (Engel, Claessens, & Finch, 2013). This dilemma is the logical outcome of the heterogeneity in mathematics skills seen in early childhood, which continues through elementary school and beyond. Hence for this study, we chose second grade as an important grade for improving mathematics achievement and for testing the causal influence of children's individual differences on mathematics achievement (i.e., child X instruction interactions).

1.1.1. *Theoretical framework*

In addition to early mathematics skills, accumulating research suggests that other child characteristics including cognitive, linguistic, and social-emotional processes, may interact reciprocally and synergistically with instruction to impact how students respond to instruction (Connor et al., 2016). Based on bio-ecological theories (Bronfenbrenner & Morris, 2006) and dynamic systems (Yoshikawa & Hsueh, 2001), the lattice model (Connor, 2016) has been applied to literacy achievement. We have adapted this model for mathematics achievement as depicted in Fig. 1. Although there are important distinctions between mathematics and reading skills, the rationale for applying this framework to mathematics emerges from evidence for shared variance among early literacy and numeracy (Davidse, De Jong, & Bus, 2014), from findings that select literacy measures predict later mathematics achievement (Purpura, Hume, Sims, & Lonigan, 2011), and for well documented comorbidity of mathematics and reading disorders (Willcutt & Pennington, 2000). Applying the lattice framework to studies of early mathematics interventions may shed additional light on shared and non-shared aspects of early learning in mathematics and reading. This also justifies including other sources of influence on children's mathematics development including gender, language skills, and socio-economic status in our models.

In the lattice model, proficient mathematics achievement relies on developing mathematics-specific processes, such as numerical processing (Mazzocco, Feigenson, & Halberda, 2011a, 2011b), other numeracy skills (Geary et al., 2013), number knowledge (Purpura et al., 2013), linguistic skills (LeFevre et al., 2010; Purpura et al., 2011), and social-emotional skills (Jones, Brown, & Aber, 2011; Rimm-Kaufman, Curby, Grimm, Nathanson, & Brock, 2009). These skills may interact with any one of a wide range of factors but particularly with instruction. For example, whereas we would not expect gender differences between boys and girls based on biological, linguistic, cognitive, or math-specific processes; it may be the case, that social-emotional attitudes and anxiety might indirectly contribute to gender differences (Beilock, Gunderson, Ramirez, & Levine, 2010). This may be one reason that gender differences do not emerge consistently on measures of the early and later numeracy and other basic mathematics skills (e.g., Geary et al., 2013; Halberda et al., 2008; Purpura et al., 2013) – the influence of gender on mathematics is indirect.

The key to our conceptual framework is that it is dynamic. That is, children bring a constellation of skills, aptitudes, attitudes, and beliefs to the process of learning mathematics and to the instructional environment. This means that what is effective mathematics

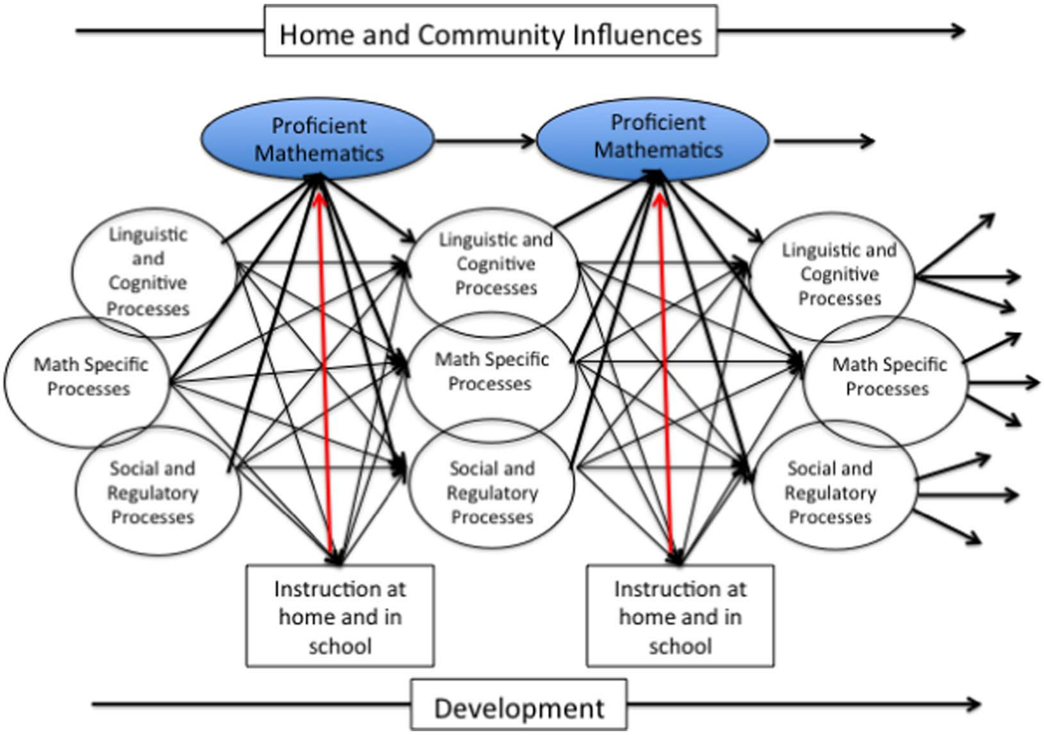


Fig. 1. Adaptation of the Lattice model for mathematics achievement. Multiple sources of influence including both proximal (e.g., child characteristics, instruction) and more distal sources of influence (e.g., school and home) that interact reciprocally and synergistically to predict global and specific aspects of students' mathematics development over time.

instruction for one child might be less effective for a child with a different constellation of skills, aptitudes, and attitudes – and that personalizing instruction, taking into account potential child X instruction interactions, should contribute to stronger mathematics achievement for children overall. Hence in this study, in addition to a range of mathematics processes, we consider students' gender, beginning of grade vocabulary knowledge (a linguistic process), and school-level poverty (as measured by the percentage of children qualifying for the US National School Lunch Program (NSLP), a widely-used indicator of family poverty). It was beyond the scope of this study to specifically test the cognitive and social-emotional aspects of the model.

1.2. Child X instruction interactions and mathematics-specific processes

In the lattice model, the instructional environment is an important source of influence on students' mathematics achievement; we would expect that student achievement would be greater when these environments are aligned with students' educational needs. In light of individual differences, children with weaker or stronger mathematics skills may not receive instruction optimally aligned with their potential achievement levels – that is, there may be child-characteristic-by-instruction (CXI) interactions. Also called skill-by-treatment interactions (Burns, Codding, Boice, & Lukito, 2010), CXI interaction effects are recognized in reading (Connor et al., 2013) but have not been as widely or as explicitly tested in elementary mathematics although there are notable exceptions (Burns et al., 2010). For example, teacher-led mathematics instruction in first grade was generally more effective for students with difficulties in mathematics whereas child-centered instruction was more effective for students with more typical mathematics skills (Morgan, Farkas, & Maczuga, 2015). Plus, there is research on personalized mathematics instruction using technology; for example, cognitive tutor technology, which is individualized based on students' performance, has been effective (Corbett, Koedinger, & Hadley, 2001; VanLehn, 2011).

Other studies of early mathematics instructional content have found less evidence of instruction aligned to students' learning needs. For example, one study found that teachers spent disproportionate time on concepts that fell either above or below a students' current level of mathematical understanding. Specifically, Engel et al. (2013) found that the kindergarten teachers in their study reported devoting most of their mathematics instructional time to counting and shape identification; but most kindergartners already had this knowledge when they entered school. Not surprisingly, Engel and colleagues also found that the children with the lowest levels of math skills were the only ones to show gains in mathematics achievement, whereas most of the other children made greater mathematics achievement gains when they were exposed to mathematical content that went beyond this basic level. Focus on basic skills continues to consume instructional time in the early grades, at least in the United States (Li, Chi, DeBey, & Baroody, 2015).

Drawing broadly from the literature on personalized or individualized instruction, (Connor, Morrison, Fishman, Schatschneider, & Underwood, 2007; Connor, Morrison, Fishman, et al., 2011; Connor, Morrison, Schatschneider, et al., 2011), an essential feature of individualized instruction is targeting instruction based on children's current performance on component skills (e.g., in reading: decoding and comprehension). A components-based approach should also work for early mathematics (e.g., Burns et al., 2010). As demonstrated in an early randomized controlled study on this approach, first graders with stronger incoming fall mathematics skills demonstrated greater mathematics gains when they were in randomly assigned classrooms where they were provided with supplementary peer-assisted learning opportunities (Fuchs et al., 1997). However, children with weaker initial skills made greater gains when they were in the control condition, which did not include peer-assisted learning opportunities (Mazzocco, Crowe, Calhoon, & Connor, in preparation). These CXI interactions are among the first to be explicitly documented for mathematics achievement (Clements, Sarama, Wolfe, & Spitler, 2013; Morgan et al., 2015).

1.2.1. *Individualizing student instruction in mathematics (ISI-Math)*

As noted previously, despite documented individual child differences in mathematics achievement, research shows that elementary mathematics instruction is rarely designed to accommodate these differences (Arnup, Cheree, John, & Louise, 2013). For example, many widely-used mathematics curricula (e.g., Saxon Math) are designed for whole-class implementation. At the same time, numerous advocates for the use of differentiated mathematics instruction recognize its effectiveness for children at various levels, from students with disabilities to those gifted in mathematics (Dee, 2010; Edwards, 2006). Despite such recommendations, many pre-service and practicing teachers are ill-prepared to differentiate instruction in their classrooms (Dee, 2010; Edwards, 2006). Given the reported absence of relevant coursework and training on children's mathematical thinking for pre-service elementary teachers (Ma, 1999), many in-service teachers appear to lack an understanding of the strategies and techniques that emphasize student development and learning that can be used to individualize instruction. Importantly, teachers' knowledge of children's mathematical learning is predictive of mathematics achievement gains among first and third graders (Hill, Rowan, & Ball, 2005). Hence, in order for an early mathematics instructional program to be effective, it should include professional development designed to support general education classroom teachers' gains in understanding children's mathematics trajectories (Turner et al., 2012).

1.2.2. *Purpose of the study*

The following two research questions guided the design and implementation of this study.

1. To what extent are CXI interactions causally implicated in early mathematics achievement? That is, if we design an individualized mathematics program, ISI Math, using assessment to guide instruction, will individualized mathematics instruction be more effective than an alternative treatment that is not individualized?

In the present study, to begin to test whether CXI interactions were causally implicated in early mathematics achievement, we conducted the cluster-randomized controlled study described here. The purpose of this study was to examine whether 2nd grade mathematics instruction that takes into account individual student differences in mathematics component skills was more effective than more typical whole group mathematics instruction provided to students in the control ISI-Reading classrooms. To test whether individualizing math instruction was generally effective, we compared students' mathematics outcomes for students in treatment (ISI-Math) vs. control (ISI-Reading) classrooms hypothesizing that mathematics instruction individualized based on students' assessed skills would be more effective in improving student outcomes than instruction provided classroom-wide to all students, regardless of skills.

2. To what extent does the ISI-Math intervention change the effect that child level gender, initial math skill, and vocabulary and school level poverty have on mathematical growth?
   The lattice model would support an association between mathematics and language (e.g., vocabulary). Thus, since we designed ISI-Math as an individualized assessment-based intervention, we hypothesized that finding CXI interactions would suggest that we had *not* succeeded in identifying important aspects of mathematics instruction that aligned with students' skills and abilities. Rather, other dimensions of instruction might be important to consider. Finding no CXI interactions would suggest we were on the right track in identifying child characteristics that interacted with mathematics instruction, and that our developmental model was supported. We included vocabulary skills as a proxy for the broader construct of linguistic skills, following the lattice model, and because, as noted previously, there is evidence of shared variance in language and mathematics skills. As noted previously, we also included gender X treatment interaction effects because there is some (albeit inconsistent) evidence that girls may have weaker math skills than do boys. Finally, we included the contextual characteristic of school-wide poverty levels. Again, if ISI-Math does consider child characteristics that impact learning, it should be effective regardless of student and school characteristics.

2. Methods

The study was conducted in one district in North Florida as part of a longitudinal study examining cumulative effects of effective instruction in reading. Second grade teachers (n = 32, 5 schools) and their students (n = 370) were randomly assigned within schools to one of two conditions – an individualized mathematics intervention (ISI-Math, n = 17 teachers), or an individualized reading intervention used as a control condition (ISI-Reading, Connor et al., 2013, n = 15 teachers). ISI-reading served as an appropriate control for this study inasmuch as we could provide exactly the same PD protocol albeit with different content aims - mathematics and reading. Most of the students had participated in the first grade study and were recruited then. This means that

Table 1
Descriptive statistics for mathematic assessments administered.

Measure	Time	Control		ISI-Math		Total	
		Mean	SD	Mean	SD	Mean	SD
Math Fluency SS	Fall	104	12.7	107	12.8	106	12.8
Math Fluency W	Fall	490	3.5	491	4.07	491	3.9
Math Fluency SS	Spring	106	13.7	110	15.3	108	14.7
Math Fluency W	Spring	493	4.5	495	5.4	494	5.1
KeyMath PR	Fall	39	25.3	41	25.8	41	25.6
KeyMath GSV	Fall	169	10.2	171	10.2	170	10.2
KeyMath PR	Spring	50	26.4	54	28.3	52	27.5
KeyMath GSV	Spring	182	11.1	185	11.6	183	11.5

Note. SS = standard score (mean = 100, SD = 15); PR = percentile rank (mean = 50, SD = 25); W = W score (SD = 15); GSV = growth scale value (SD = 10). Fall n = 212 ISI-Math and 178 ISI-Reading; Spring n = 205 ISI-Math and 165 ISI-Reading; Attrition = 5%.

approximately half of the children had participated in the ISI-Reading intervention and half in Math PALS the previous year (Fuchs et al., 1997). We followed them into second grade and recruited their teachers by visiting the school, discussing the study, and inviting them to participate. We then recruited classmates who had not participated in the first grade study and obtained parental consent for the remaining two years of the study. At the completion of the second grade study, we followed students into third grade, and recruited their teachers. Results of the ISI-Reading study are reported elsewhere (Connor et al., 2013). The purpose of the present study was to examine the impact of ISI-Math on second grade students' mathematic achievement gains. ISI-Math was specifically developed to test whether CXI interactions are causally implicated in second graders' varying mathematics achievement.

Random assignment of teachers was conducted immediately after the start of the school year once students were assigned to classrooms. School percentages of students participating in the National School Lunch Program (NSLP) ranged from 33% to 59% (M = 45%). All of the teachers were women and White. On average, they had 6.7 years of experience teaching second grade and 14.5 years of experience teaching overall. All had at least a bachelor's degree, 24% had an MS or MA degree, and 12% had a MEd or MAT degree; these demographic descriptors did not vary by assigned condition.

2.1. Student participants

Of the 370 2nd graders who participated in the study, 192 (52%) were girls, and 209 (56%) were in ISI-Math classrooms. About 84% were White, 5% were African American, 5% were multiracial, and the rest belonged to other ethnicities, with an equal distribution across conditions. Student attrition was 5% (see Table 1). No schools or teachers withdrew from the study.

2.1.1. Professional development

Teachers in both the ISI-Reading and ISI-Math conditions received the same professional development (PD) protocol except for content. This included a half-day workshop at the beginning (August) and middle (January) of the school year, monthly communities-of-practice meetings at their school (Bos, Mather, Narr, & Babur, 1999), and hour-long bi-weekly instructional coaching visits during mathematics (or reading) instruction (Carlisle, Kelcey, Rowan, & Phelps, 2011). PD was provided by one of several certified research teachers who were members of the research team, called research partners. The workshops focused on defining ISI-Math, introduced the *Concepts and Applied Skills Assessment* (CASA, described subsequently and in Appendix A), and gave teachers opportunities to ask questions and meet with teachers at other schools and grades. The workshop in January focused on the results of standardized assessments and how to interpret them.

PD during the grade-level and school-based communities of practice (Bos et al., 1999), also known as professional learning communities, focused on using the ISI-Math assessment, *Concepts and Applied Skills Assessment* (CASA, described later, see Appendix A for examples of items), to group children into flexible learning groups and to discuss effective mathematics instruction using best practices (National Mathematics Advisory Panel, 2008). For example, in grade level teams during the community of practice, the teachers would review the CASA results for the students in their classroom, the groups they were in currently, and then develop new groups aligned with students' changing performance on the CASA. If teachers had access to other data (e.g., the Math Fluency assessments conducted in January), they reviewed those data as well and then developed weekly lesson plans in collaboration with their research partner. The research partner also provided printed learning activities to supplement those available in the mathematics core curriculum. The aim was to establish groups of children with similar learning needs, based on CASA results, and then provide mathematics learning opportunities that aligned with the assessment data. Other PD topics included classroom management and setting up stations so that teachers could work with the learning groups while the other children worked independently or with peers on assessment-aligned mathematics learning activities; using research to inform practice; and improving knowledge about mathematics and best practices. The aim in these communities of practice was to turn the responsibility for using data and designing instruction back to the teachers using a release of responsibility model (Pearson & Gallagher, 1983).

In-classroom PD, which occurred during the dedicated time devoted to mathematics instruction, focused on problem solving, modeling best practices, assisting with groups, and classroom management. Following the in-class time, teachers would meet with their research partner individually to discuss challenges and specific questions they might have.

Teachers in the ISI-Math condition typically provided whole-class core curriculum mathematics instruction followed by ISI-Math individualized small group time (i.e., stations) for about 20 min. Teachers in the control ISI-Reading condition continued to provide their typical math instruction based on classroom observation. ISI-Math materials were developed by researchers and by adapting materials developed for MATH-Pals (Fuchs et al., 1997).

2.2. Implementing ISI-Math

ISI-Math was designed as a general education classroom intervention to be provided to all students, not only those struggling with mathematics. This is because individualized instruction does not assume basic skills remediation unless remediation is indicated by students' performance levels. The protocol is designed to promote the use of more challenging content matched to students' skills. Therefore, for the randomized controlled trial (RCT), ISI-Math was implemented during the dedicated block of time devoted to mathematics instruction and utilized *stations*. During station time, the classroom teacher met with small groups of students at the teacher table while the rest of the students worked individually or with peers completing appropriate learning activities. Again, using the CASA, which was developed specifically for ISI-Math, students were grouped according to learning needs into flexible learning groups. ISI-Math activities, which included newly developed items as well as items adapted from the Math-PALS program (http://kc.vanderbilt.edu/pals/math.html, examples available upon request from the first author), were designed to follow a developmental scope and sequence of learning progressions (National Council of Teachers of Mathematics (NCTM), 2013) aligned with skills evaluated using the CASA. Thus, small-group instructional activities were matched with students' areas of strengths and weaknesses.

Teachers provided instruction across topics – numeracy, measurement, geometry, algebra, etc., based on results of the CASA, rather than on the current topic of their core mathematics curriculum (i.e., Saxon Math). This meant that during station time, students were mastering different skills based on where their mathematic skills fell on the CASA, which was organized according to generally accepted mathematics topics continuum (National Council of Teachers of Mathematics (NCTM), 2013). Learning activities included both practice-based and problem-based activities.

Generally, practice-based activities are more closely aligned with approaches that involve explicit instruction of mathematics procedures followed by repeated student practice (during the time they are not meeting with the teacher at the teacher station). In contrast, problem-based activities are more closely aligned with approaches to mathematics instruction in which students explore a task through reasoning and logic to discover mathematical ideas or concepts (National Council of Teachers of Mathematics (NCTM), 2013), usually working with the support of other students (Morgan et al., 2015). ISI-Math incorporated both types of activities.

During ISI-Math stations, instruction was provided in highly interactive small flexible learning groups with the teacher, where students with similar skills and learning needs (based on the CASA) were grouped together. As students' skills improved, new content was introduced and students were re-grouped. Regrouping could occur as frequently as once a month and always after the CASA was administered. Teachers typically had four groups but since this was at the teachers' discretion, some had five smaller groups or met with some children individually. Teachers met with each group at least twice a week for about 20 min with more time per week provided for students with weaker skills. According to CXI interaction research, in general, students with weaker skills require more time in small groups with their teacher than do children with stronger skills (Connor, Morrison, Fishman, et al., 2011; Connor, Morrison, Schatschneider, et al., 2011; Morgan et al., 2015). While teachers met with the small groups, the other students worked independently or with their peers using previously introduced ISI-Math learning activities, which were matched to their specific strengths and weaknesses, and complemented the activities implemented at the teacher station. Importantly, as students' learning needs changed, so too did grouping configurations. Literacy research indicates that such teacher-managed dynamic small group instruction is four times more effective than is instruction on the same topic delivered to the entire class, and appears to increase student engagement (Connor, 2011).

To determine small group membership, students were administered the CASA progress monitoring assessments developed for the ISI-Math intervention 4 times per year – and, as discussed, were reviewed at the communities of practice. Areas of need in foundational skills were the first targets of instruction, followed hierarchically by skills that built upon mastered skills. For example, number sense skills were prioritized above measurement skills since many measurement concepts depend on an understanding of number, addition, and place value.

As described previously, teachers received professional development to support their implementation of ISI-Math and were taught the following evidence-based guidelines:

1. Students struggling with math generally make greater gains when they spend more time in explicit and systematic mathematics instruction conducted in small groups with their teacher (Hiebert, Gallimore, & Stigler, 2002; Morgan et al., 2015).
2. More skilled students make greater gains when they are challenged and allowed to work independently and with peer-assisted learning interventions (Baker, Gersten, & Lee, 2002; Leung, 2015; Morgan et al., 2015).
3. Students with typical math skills may not be getting the right balance of systematic and explicit instruction versus challenging work (Doyle, 1988).
4. When students are working independently or with peers;
   a. The work should be meaningful and appropriately challenging in order to maintain student engagement (Henningsen & Stein, 1997).
   b. Students should generally be on-task (Stallings, 1980).
   c. Students who are stronger in mathematics are more likely to make greater gains when their student-managed times are more conceptual and applied; they generally make weaker gains when the activities involve practice of mathematics skills and

concepts they have already mastered (Rogers, 2000).

d. Students who are struggling with mathematics generally make greater gains when they have enough time to practice the conceptual and applied mathematics skills they are learning (Morgan et al., 2015). The optimal amount of individual and group work with peers increases as the year progresses and students' mathematics skills and ability to apply these skills improve (i.e., gradual release of responsibility to students).

All teachers (across both conditions) used Saxon Math (http://saxonpublishers.hmhco.com/en/sxnm_home.htm) as their core mathematics curriculum, which was selected by their school district. Generally, Saxon Math emphasizes a traditional approach to teaching and learning mathematics that highlights direct, explicit instruction with less emphasis on real world connections and student thinking. It provides frequent opportunities for students to practice number and operations in a drill-based manner (Sood & Jitendra, 2007) and is designed to be used whole class for all students regardless of ability.

2.2.1. Child assessments

The CASA measures students' knowledge of mathematics concepts and their ability to apply those skills. Items range in difficulty from kindergarten through 5th grade levels and cover the basic mathematical skills where mastery is expected for each grade. The 93 items in the CASA are divided into sections by mathematical skill areas. These skills include: numeration, measurement and data analysis, fractions, geometry, and word problems. Each section includes items that increase in difficulty to determine students' proficiency and instructional levels. Subscales and subscores are also available for each section. Reliability for the total score with this sample was excellent (alpha = 0.92).

Reports of the CASA were shared with teachers during the monthly community of practice (COP) that immediately followed administration, with four administrations over the school year. Reports included results for each student in table form, focusing on skills mastered and not mastered (see Fig. 2), and, graphically, showing gains in scores over the school year. For some students, particularly those who were not making expected progress, the teachers and research partners would review the entire test. The research partners and teachers together decided how to group students and which aspects of mathematics to focus on during stations. As the school year progressed, teachers became more independent in their decision-making as the research partners turned over decision-making responsibilities to teachers using a release of responsibility model (Pearson & Gallagher, 1983). The CASA was specifically designed to facilitate this kind of instructional decision-making.

Children's fall and spring mathematics skills were assessed using two widely adopted standardized achievement measures, the *Woodcock-Johnson III* (Woodcock, McGrew, & Mather, 2001) Math Fluency subtest and the *KeyMath3* (Connolly, 2007). Both of these were used as the outcome measures for this study because they reflect related but independent mathematics skills (Petrill et al., 2012) that predict later achievement.

The WJ-III Math Fluency subtest asks children to solve increasingly more difficult addition, subtraction, multiplication and division problems using paper and pencil. The task is overtly timed and children are asked to complete as many problems as possible, as quickly as possible. Performance outcomes are based on the number of correctly completed problems, which are converted to W scores based on grade at testing. W scores, which have equal intervals, were used in the analyses; standard scores ($M = 100$, $SD = 15$), which control for age/grade, are presented to interpret the findings. Reliability for the measure was excellent (alpha = 0.90).

The KeyMath-Third Edition (Connolly, 2007) includes multiple subtests designed to assess skills within greater mathematics domains of basic concepts (subtests include numeration, algebra, geometry, measurement, data analyses & probability), operations (subtests include mental computation, estimation, addition, subtraction, multiplication, and division), and applications (which includes one subtest, problem solving). We used the total growth scale value (GSV) as our outcome. Percentile ranks ($M = 50$, $SD = 25$) based on age-references norms were used to interpret the scores.

We also assessed children's vocabulary skills using the *WJ-III Picture Identification* task as a proxy for linguistic skills in the lattice model. In this individually administered task, children name pictures of increasingly unfamiliar objects and concepts. As with the Math Fluency assessment, W scores were used in the analyses and reliability was acceptable for this age range ($r = 0.81$). Unfortunately, we did not have direct measures of children's cognitive, social-emotional or behavioral skills and so were not able to test these aspects of the lattice model.

2.2.2. Fidelity of implementation

To monitor fidelity of implementation, classroom observations were conducted live, as well as video-taped, during both mathematics and literacy instruction, three times per year in the fall, winter, and early spring. Each observation lasted approximately 45–60 min and was scheduled at the teachers' convenience. These observations were then coded using a brief version of the ISI Observation System (Connor et al., 2009). Results of these observations revealed that ISI-Math was delivered as intended although there was notable between classroom variability. Observations showed that rarely did ISI-Reading teachers use ISI-Math techniques during mathematics instruction, keeping in mind that CASA results were not shared with the ISI-Reading control group teachers.

The *Individualizing Student Instruction Fidelity Scale* (short form), which includes a rubric for classroom implementation of individualized instruction, teacher's organization, and their warmth/responsiveness and disciplinary affect, on a scale of 1 (poor) to 6 (exemplary) was completed at the time of the observations by trained research assistants. Inter-rater agreement was computed for approximately 10% of the observations at each time point, and was acceptable, with kappa ranging from 0.73 to 0.82.

Teachers in classrooms that were rated highly on the Individualization dimensions fully implemented multiple and flexible student grouping configurations and regrouping of students based on formal and informal assessment data. The content of literacy/

Teacher	Child Last Name	Child First Name	Add 1x1	Sub 1x1	Add 2x1	Sub 2x1	Add 2x2	Sub 2x2	Add 3x3	Sub 3x3	mult 1x1	nult 2x1, 2x	nult 3x1, 4x	Div	group
	a		41	0											1
	b		41	0											1
	c		41	7											1
	d		41	50											1
	e		76	0											2
	f		76	29											2
	g		76	43											2
	h		82	36											3
	i		88	0											3
	j		88	36											3
	k		88	71											3
	l														
			82	93	33	50	0								4
	m		82	100	44	100	11								4
	n		94	100	33	100	22								4
	o		94	100	78	100	67								4
	p		100	86	22	50	33								4
	q		100	100	89	100	56								4

Teacher	ild Last Na	ild First Na	Add 1x1	Sub 1x1	Add 2x1	Sub 2x1	Add 2x2	Sub 2x2	Add 3x3	Sub 3x3	mult 1x1	nult 2x1, 2x	nult 3x1, 4x	Div		group
	a		47	86	44											1
	b		53	14	11											1
	c		82	50	33	100	33									1
	d		82	71	33	50	44									1
	e		94	79	33	0	78									2
	f		94	71	89	100	44									3
	g		94	57	89	100	44									2
	h		94	71	33	50	67									2
	i		94	86	0	50	22									3
	j		88	93	33	0	56	25								2
	k		100	86	56	50	89	0								4
	l		100	93	33	50	89	100								4
	m		100	100	44	50	33	75								4
	n		94	86	44	100	33	25	0	0						3
	o		94	93	78	100	22	0	0	0						3
	p		100	86	33	100	89	75	100	33						4
	q		100	100	100	100	56	100	50	33						4
	r		88	93	100	100	89	100	50	67						4
	100-80															
	50-79															
	0-49															

Fig. 2. Two examples of how students' CASA results (blinded) were displayed graphically for teachers to aid in flexible learning group development and focus of learning activities by group. The top chart shows children at the beginning of the year; the bottom chart shows children later in the school year. Squares that are in green (light gray) indicates mastery of at least 80% in that area; yellow (medium gray) indicates mastery of at least 50%, and red (black) indicates mastery below 50%. Students were grouped based on skills where mastery fell below 80% with focus on earlier skills (e.g., Adding 1 × 1) that were coded red (i.e., black) or yellow (medium gray).
Note that for the most part, children mastered skills in developmental order but not always. In this case, a specific child might be put in a group with other children working on a specific skill even though the child had mastered more difficult skills (e.g., child L in bottom chart). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

math instruction was differentiated. The entire language arts/math block was spent in meaningful literacy/math activities. For classrooms that were rated highly on the organization dimensions, observers reported that the classroom was well organized and instruction was clear and well organized. Classroom routine was evident. Transitions were efficient. Classrooms that were rated highly on the warmth/responsivity/control/discipline dimension had teachers who were observed to be authoritative but never punitive. Teachers effectively selected and incorporated students' responses, ideas, examples, and experiences into the lesson. The classroom consistently offered a positive learning environment with clear expectations for students' behavior as a member of the learning community.

Overall quality of the classroom learning environment was generally good (average scores between 3 and 4 for Individualizing and Organization, and between 4 and 5 on Warmth and Discipline). We used Multivariate ANOVA (SPSS v22, General Linear Modeling) to compare ISI-Math and ISI-Reading classrooms across the three dimensions for fall, winter and spring. Results indicated no significant differences in the fall [Wilks' Lambda = 0.928, $F(3, 28) = 0.720$, $p = 0.548$]. In the winter and the spring ISI-Math teachers were rated significantly higher than ISI-Reading [Winter Wilks' Lambda = 0.669, $F(3, 28) = 4.612$, $p = 0.010$; Spring Wilks' Lambda = 0.508, $F(3, 29) = 9.357$]. Post-hoc results indicated that ISI-Math teachers received significantly higher ratings for Individualizing in Winter and Spring compared to the fall observation and compared to ISI-Reading teachers. Organization and Warmth/Discipline ratings were stable over time and did not differ by group.

3. Results

In the fall, at the beginning of the study, students in ISI-Math and ISI-Reading classrooms were achieving Math Fluency scores at or above grade expectations but KeyMath scores were somewhat below grade expectations (see Table 1). Students made greater than expected gains from fall to spring, with Math Fluency standard scores above the expected mean of 100 and spring KeyMath percentile ranks at grade expectations. However, within the sample, the ranges of achievement scores were large with spring Math Fluency standard scores from 74 to 164 and KeyMath percentile ranks from 1 to 98. Children's vocabulary scores also fell well within typical ranges with standard score means of 103 ($SD = 9.4$) in the fall and 103 ($SD = 9.3$) in the spring. All assessments were normally distributed with metrics of skewness and kurtosis close to zero (i.e., no value > 1 or less than -1). Visual inspection of the data as well as Q-Q plots confirmed normal distribution.

3.1. Research question 1: to what extent are CXI interactions causally implicated in early mathematics achievement?

To assess whether individualizing mathematics instruction based on children mathematics skills was more effective than the control, we used hierarchical linear modeling (HLM, Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2004) because students were nested in classrooms, with the treatment variable entered at the classroom level, which was the level of randomization. ISI-Math was coded 1; ISI-Reading was coded 0. In all reported models, we used restricted maximum likelihood; running the models using full maximum likelihood yielded the same results (Peugh, 2010). All continuous variables were group-mean centered, which provides more unbiased estimates when investigating skill X treatment interaction effects. Residuals from all HLM analyses were normally distributed.

We first ruled out differences in initial status for the treatment and control students for mathematics achievement, vocabulary, and gender (Stoet & Geary, 2015). HLM analyses revealed no significant differences by condition for fall KeyMath or Vocabulary. Girls scored significantly lower on fall KeyMath than did boys (2.4 points lower, see Table 2), so gender (girl = 1; boy = 0) was included in all models. There were group differences in fall Math Fluency, favoring the ISI-Math group and so all models controlled for fall Math Fluency (group mean centered).

Using unconditional models, we computed intraclass correlations (ICC), which reflect the proportion of between classrooms variance explained. The classroom level variance was significantly > 0 ($p < 0.05$). The ICC for spring Math Fluency was 0.195 and for KeyMath was 0.063.

End of second grade results revealed a significant main effect of treatment condition: students in the ISI-Math classrooms achieved significantly higher scores on Math Fluency ($d = 0.60$, see Table 3 & Fig. 3) and for KeyMath ($d = 0.41$, see Table 4 & Fig. 3), taking into account gender and fall Math Fluency, compared to students in ISI-Reading classrooms.

3.2. Research question 2: to what extent does the ISI-Math intervention change the effect that child-level gender, initial math skill, and vocabulary, and school-level poverty have on mathematical growth?

Again, we hypothesized that because ISI-Math was designed to be individualized and assessment-driven, there would be no significant skill X treatment interactions and these hypotheses were supported. There was no effect of gender; nor was there a gender X treatment interaction. We found no significant gender differences for fall or spring Math Fluency, and no significant between-school differences for school NSLP percentages for KeyMath or Math Fluency. School NSLP was entered at the classroom level but trimmed from the final models to preserve parsimony.

We then examined whether fall Vocabulary was associated with spring KeyMath and Math Fluency outcomes. Fall Vocabulary did not predict spring Math fluency (coefficient $= -0.002$, $p = 0.92$) nor were there child fall vocabulary X ISI-Math interaction effects on spring Math Fluency ($p = 0.50$) and so the variable was trimmed from the model to preserve parsimony. Fall vocabulary did significantly predict spring KeyMath (see Table 5) but there was not a significant fall vocabulary X ISI-Math, which was trimmed to preserve parsimony.

Table 2
Effect of intervention condition, gender, and school poverty on fall KeyMath.

Fixed effect	Coefficient	Standard error	*t*-Ratio	Approx. *d.f.*	*p*-Value
For intercept, β_0					
Fitted mean, γ_{00}	171.871	2.552	67.354	29	< 0.001
ISI-Math, γ_{01}	1.918	1.084	1.769	29	0.087
SCHNSLP, γ_{02}	− 0.029	0.061	− 0.474	29	0.639
For MFW slope, β_1					
Effect, γ_{10}	1.884	0.181	10.392	334	< 0.001
ISI-Math X MFW, γ_{11}	− 0.559	0.245	− 2.281	334	0.023
For girl slope, β_2					
Effect, γ_{20}	− 2.447	0.784	− 3.122	334	0.002

Final estimation of variance components

Random effect	Standard deviation	Variance component	*d.f.*	χ^2	*p*-Value
Intercept, u_0	1.852	3.431	29	44.921	0.030
Level-1, *r*	8.272	68.431			

Note. ISI-Math = 1, Control = 0; MFW = fall Math Fluency W score group mean centered; Girl = 1, Boy = 0. SCHNSLP = School percentage of students qualifying for the National School Lunch Program. Deviance = 2615.76.
Mixed model
Fall Key Math $GSV_{ij} = \gamma_{00} + \gamma_{01} * \mathit{ISI\text{-}Math}_j + \gamma_{02} * \mathit{SCHNSLP}_j$
$+ \gamma_{10} * \mathit{MFW}_{ij} + \gamma_{11} * \mathit{ISI\text{-}Math}_j * \mathit{G2MFW}_{ij}$
$+ \gamma_{20} * \mathit{Girl}_{ij}$
$+ u_{0j} + r_{ij}$.

Table 3
Effect of intervention condition, gender, and fall Math Fluency on spring Math Fluency W score.

Fixed effect	Coefficient	Standard error	*t*-Ratio	Approx. *d.f.*	*p*-Value
For intercept, β_0					
Fitted mean, γ_{00}	493.366	0.548	899.101	30	< 0.001
ISI-Math, γ_{01}	1.849	0.845	2.186	30	0.037
For MFW slope, β_1					
Effect, γ_{10}	0.872	0.056	15.530	338	< 0.001
For Girl slope, β_2					
Effect, γ_{20}	− 0.198	0.316	− 0.629	338	0.530

Random effect	Standard deviation	Variance component	*d.f.*	χ^2	*p*-Value
Intercept, u_0	2.322	5.392	30	238.439	< 0.001
Level-1, *r*	3.102	9.623			

Note. 42% of variance in spring Math Fluency explained. ISI-Math = 1, Control = 0; MFW = fall Math Fluency W score group mean centered; Girl = 1, Boy = 0. Deviance = 1959.158.
Mixed model
Spring Math Fluency $W_{ij} = \gamma_{00} + \gamma_{01} * \text{ISI-Math}_j + \gamma_{10} * \text{MFW}_{ij} + \gamma_{20} * \text{Girl}_{ij} + u_{0j} + r_{ij}$.

Math Fluency skills in the fall significantly and positively predicted students' mathematics scores on both of our outcomes – Math Fluency and KeyMath. However, there were no fall Math Fluency X treatment interaction effects on either outcomes. That is, ISI-Math was effective regardless of students' incoming mathematics skill. Of note, the ISI-Math and ISI-Reading control group students showed gains in Math Fluency relative to their own fall scores, possibly reflecting the benefits afforded by a curriculum that places emphasis on fluency practice (Fuchs, Seethaler, Fletcher, & Hamlett, 2009), which was a component of the schools' curriculum, Saxon Math. Findings also speak to the specificity of learning when comparing fitted mean scores for Math Fluency and KeyMath (see Fig. 3). On average, students in both groups demonstrated stronger achievement on Math Fluency (a central component of Saxon Math Curriculum) than on KeyMath, which focuses on a broader range of mathematic skills and concepts. Unfortunately, there were too few items within these subdomains to allow testing for effects on specific mathematic content areas.

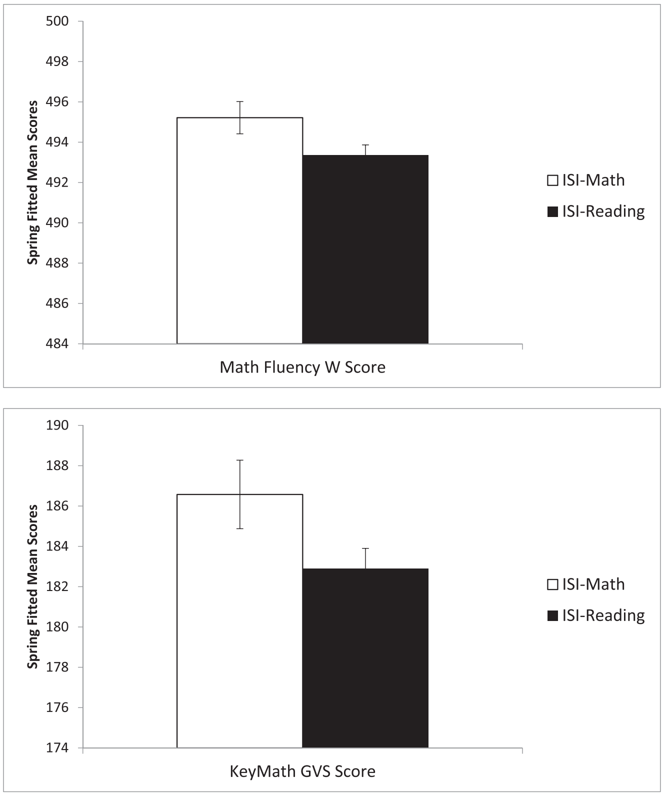


Fig. 3. Comparing mean spring scores controlling for fall score and gender for students in the ISI-Math and ISI-Reading control condition on standardized measures of mathematics. Range on y-axes represents 1 SD. Error bars are standard errors.

Table 4
Effect of intervention condition, gender, and fall Math Fluency on spring KeyMath scale value (GVS) scores.

Fixed effect	Coefficient	Standard error	*t*-Ratio	Approx. *d.f.*	*p*-Value
For intercept, β_0					
Fitted mean, γ_{00}	182.899	1.052	173.812	30	< 0.001
ISI-Math, γ_{01}	3.673	1.730	2.123	30	0.042
For MFW slope, β_1					
Effect, γ_{10}	1.879	0.220	8.488	334	< 0.001
ISI-Math X MFW, γ_{11}	− 0.425	0.274	− 1.550	334	0.122
For girl slope, β_2					
Effect, γ_{20}	− 1.375	1.411	− 0.975	334	0.330
ISI-Math X girl, γ_{21}	− 2.623	1.915	− 1.369	334	0.172

Random effect	Standard deviation	Variance component	*d.f.*	χ^2	*p*-Value
Intercept, u_0	3.826	14.637	30	94.65528	< 0.001
Level-1, *r*	8.881	78.867			

Note. 38% of variance in spring KeyMath scores explained. ISI-Math = 1, Control = 0; MFW = fall Math Fluency W score group mean centered; Girl = 1, Boy = 0. Deviance = 2689.397.
Mixed model
$Spring\ Key\ Math_{ij} = \gamma_{00} + \gamma_{01} * ISI\text{-}Math + \gamma_{10} * MFW_{ij} + \gamma_{11} * ISI\text{-}Math_j * MFW_{ij}$
$+ \gamma_{20} * Girl_{ij} + \gamma_{21} * ISI\text{-}Math_j * Girl_{ij} + u_{0j} + r_{ij}.$

4. Discussion

In this study, we showed that teachers' instructional practices do affect students' mathematics achievement gains, and are more effective when individual student differences in mathematics skills are considered. This finding supports our theoretical framework, the hypotheses generated, and the causal implications of CXI, specifically skill X treatment interactions in mathematics. We found

Table 5
Effect of ISI-Math intervention condition on spring KeyMath including fall vocabulary (PVW), gender, and fall Math Fluency.

Fixed effect	Coefficient	Standard error	*t*-Ratio	Approx. *d.f.*	*p*-Value
For intercept, β_0					
Fitted mean, γ_{00}	169.027	0.859	196.672	30	< 0.001
ISI-Math, γ_{01}	2.887	1.151	2.506	30	0.018
For MFW slope, β_1					
Effect, γ_{10}	1.292	0.139	9.295	333	< 0.001
For PVW slope, β_2					
Effect, γ_{20}	0.403	0.045	8.866	333	< 0.001
For girl slope, β_3					
Effect, γ_{30}	0.341	0.989	0.345	333	0.730
ISI-Math X girl, γ_{31}	− 1.849	1.224	− 1.510	333	0.132

Random effect	Standard deviation	Variance component	*d.f.*	χ^2	*p*-Value
INTRCPT1, u_0	2.18635	4.78011	30	58.17541	0.002
Level-1, *r*	7.54614	56.94421			

Note. 53% of variability in spring KeyMath scores explained. ISI-Math = 1, Control = 0; MFW = fall Math Fluency W score group mean centered; PVW = fall Vocabulary W score group mean centered; Girl = 1, Boy = 0. Deviance = 2551.833.
Mixed model
Spring $\text{KeyMath}_{ij} = \gamma_{00} + \gamma_{01} * \textit{ISI-Math}_j$
$+ \gamma_{10} * MFW_{ij}$
$+ \gamma_{20} * PVW_{ij}$
$+ \gamma_{30} * Girl_{ij} + \gamma_{31} * \textit{ISI-Math}_j * Girl_{ij}$
$+ u_{0j} + r_{ij}$.

that providing professional development to general education teachers about how to individualize mathematics instruction dynamically, using CASA progress monitoring assessment results and flexible learning groups, led to student achievement gains above and beyond the grade appropriate gains associated with generally effective teaching and use of the Saxon Math curriculum. Moreover, we found that ISI-Math was effective regardless of students' fall mathematics skills even though such skill X treatment interaction effects have been found in several other studies (e.g., Morgan et al., 2015). Students' mathematic gains were not solely the result of additional professional development because teachers in the ISI-Reading control group also received professional development on individualized instruction, albeit in reading. Classroom observations revealed that teachers did individualize instruction more in the subject area, mathematics or reading, for which they received training (see also Connor et al., 2013), demonstrating that the professional development protocol used in this randomized controlled trial changed practices and improved student outcomes in mathematics (or reading depending on group assignment). Importantly, the professional development was designed to promote sustained support for teachers by providing continuity of feedback through communities of practice and in-classroom support, as well as workshops provided in the fall and winter.

Overall ISI-Math students' KeyMath and Math Fluency performance gains were equivalent to an additional 4-months of schooling relative to control students (see Fig. 3), which emerged over one school year (approximately 9 months). For example, at the end of second grade, students in the ISI-Reading control condition were generally achieving just below grade level expectations (mean grade equivalent [GE] score = 2.8) for KeyMath, whereas students in the ISI-Math condition performed above grade level expectations (mean GE score = 3.2) – this difference was greater for Math Fluency. Moreover, regular classroom teachers provided this more effective instruction, which reveals the importance of targeted professional development and support. These results also show that general education teachers can learn to interpret and use assessment results, such as the CASA, to inform their practice and the amounts and types of mathematics instruction they provide. Overall, they can teach mathematics more effectively. We see that changes in their classroom instructional practices can have an immediate and educationally significant impact on their students' mathematics learning.

In terms of mathematics teaching and learning, two approaches are seen most frequently in practice (Boaler, 2002). The traditional approach primarily focuses on learning mathematics with the teacher explaining procedures and the students practicing the procedures on their own. The Saxon textbooks used by the teachers and students in this study align most closely with this approach. The second primary approach, often called reform, focuses on learning mathematics through discovery. Students more often work in groups and collaborate in discovering mathematical patterns and relationships among numbers and their operations with the teacher serving as the facilitator (Sood & Jitendra, 2007). With regard to widely-used theories of instruction in mathematics, there is controversy about whether to use traditional or reform approaches (Lubienski, 2004; Reys, 2001; Schoenfeld, 2004; Van den Heuvel-Panhuizen, 2010). A reform-based approach is typically recognized as having the potential to improve students' understanding of mathematical concepts (Fosnot & Dolk, 2001; Gravemeijer, 2014; Hiebert, 1999; Hiebert et al., 2002; Schoenfeld, 1992). However, there are studies to support the integration of traditional approaches using Saxon Math (Agodini et al., 2009; Resendez & Azin, 2005; Resendez & Sridiharan, 2006) and reform-based approaches (Bryant et al., 2008; Morgan et al., 2015; Riordan & Noyce, 2001; Rivera,

2014) in relation to teaching number and operations to children in the primary grades. The results of this study suggest that both approaches are effective but the impact of each will depend on the skills students bring to the classroom.

With ISI-Math, we incorporated both approaches based on students' assessed learning needs: (1) teacher/child-managed small group instruction, which was highly interactive – a scaffolded reform approach so to speak – but also incorporated explicit instruction when it was needed (more traditional); and (2) student/peer-managed instruction which allowed students to interact with their peers in problem solving (reform) but also to practice the skills they were learning (traditional). Moreover, ISI-Math adopted a developmental approach to selecting mathematic topics (Clements et al., 2013) rather than staying wedded to the scope and sequence of the curriculum. Hence, those students who were ready to move on to more complex and sophisticated mathematics topics did so; whereas those students who were struggling received extra support so that they could move ahead more quickly once they mastered more fundamental aspects of mathematics. The adoption and integration of potentially more effective teaching practices used in ISI-Math (Agodini et al., 2009; Bryant et al., 2008) may have further supported the students' growth and mathematical understanding.

The results of this study do not support one approach over another. Rather they show that when teachers implement ISI-Math and thereby intentionally select among common teaching mathematics methods to personalize instruction to their students' current levels of mathematics skills and knowledge, based on assessment data, and thus appropriately challenge students with new topics and knowledge, their students generally make significantly greater gains over a school year.

4.1. Support for the lattice model as a conceptual framework

We specifically tested parts of the lattice model (see Fig. 1) with this study and found emerging support. First, we found no child characteristics X ISI-Math interaction effects for key predictors including fall mathematics, vocabulary, gender, and school-wide poverty. At the same time, fall math and vocabulary predicted spring KeyMath outcomes, which would be expected within the lattice model.

In our study, girls had lower KeyMath scores in the fall of second grade than did boys. This gender difference was not trivial with a marked effect size ($d = 0.29$); it was limited to the KeyMath test and was not apparent by spring (i.e., gender did not predict spring outcomes). Notably, KeyMath is an untimed, broader measure of mathematics than is Math Fluency, which is strongly influenced by regular short practice intervals (Fuchs et al., 2009). We can only speculate on the reasons for the initial gender differences. Perhaps girls' weaker fall KeyMath performance reflect different degrees of instructional attention for boys compared to girls in first grade (the year prior to this study), but this is unlikely given that the difference was present early in the school year of second grade but not later. Following the lattice model we might expect indirect effects of gender on achievement. For example, there is some evidence that mathematics achievement growth in primary school-age girls is subject to stereotype threat, to teachers underrating girls' proficiency, and teacher's own anxiety, compared to achievement growth in boys – at least in classrooms with female teachers (Beilock et al., 2010; Lubienski, 2004; Robinson-Cimpian, Lubienski, Ganley, & Copur-Gencturk, 2014). We did not measure any of these factors in the current study, given our lack of focus on gender differences per se. Importantly, there were no gender differences or gender X treatment interactions that affected the impact of ISI-Math.

4.2. Limitations

There are limitations to this study that should be noted when interpreting these results. First, as with any school-based RCT, there are external and internal sources of bias that cannot be controlled. For example, students move out of the school or district (there was 5% attrition in students from fall to spring), teachers go on leave, and external school and local policies influence teachers' fidelity to the implementation of the intervention. One such instance of the latter was that in some schools, teachers focused on test preparation rather than implementing ISI (either -Math or -Reading) in the spring just before state-wide testing. Additionally, although fidelity was generally strong, there was variability in implementation, which is clear from the significant classroom level variance (see Tables 3–5). Indeed, the conditional ICCs, which represent the variance between classrooms for the final models, ranged from 0.15 for spring KeyMath to 0.35 for Math Fluency. Also, it is possible that using an alternative treatment, ISI-Reading, which was highly similar to ISI-Math, except for the topic of instruction, may have attenuated the treatment effects. Other control group treatments would help to elucidate specific aspects of ISI-Math. For example, comparing ISI-Math with another supplementary mathematics program that is not data driven would have explicated the effect of the assessment-to-instruction model used in ISI-Math. Additionally, for the fidelity measure, we computed inter-rater reliability on 10% of the observations. Using a higher percentage of observations to determine reliability would have been preferable. Because this was a longitudinal study, a significant number of the students participated in the study as first graders either in the ISI-Reading or Math PALS condition. The experience in first grade may have influenced our results inasmuch as effects of condition were likely cumulative (Connor et al., 2013). Finally, as with any study, findings should be replicated in other grades and with other populations. In our study, we had a higher than average proportion of children who qualified for NSLP so these findings might not generalize to more affluent schools and students.

4.3. Implications and future research

Our primary findings demonstrate that CXI interactions, which represent the same construct as skill X treatment interactions (Burns et al., 2010), appear to be causally implicated in students' mathematics achievement, and that data-based decision-making to personalize mathematics instruction is possible, and effective, in the early school years. We observed achievement gains for both fluency and a variety of mathematics skills, despite a very large range of mathematics performance levels (1st to 98th percentile) in our sample, and despite the relatively brief duration of the small-group individualized instruction experienced by students (as little as 20 min, 2 times per week). Higher doses of individualized instruction may be even more effective, at least for some students. Additionally, ISI-Math, informed by the lattice model, intentionally incorporated important components of both reform and traditional mathematics educational approaches, based on students' assessed learning needs, and took into account linguistic processes (i.e., vocabulary) and school SES.

Data-based decision making, which is at the heart of ISI-Reading, has a strong foundation in research (Fuchs, 2004). Still, more research is needed to understand the optimal combinations and formats of data-informed individualized mathematics instruction, which sources of individual child differences are of greatest importance for shaping instructional plans (e.g., English language proficiency, self-regulation), and precisely how these combinations affect student achievement. This research supports more complex models of instruction, such as the lattice model. The results can also inform the development of dynamic forecasting intervention algorithms similar to those developed for kindergarten through 3rd grade reading (Al Otaiba et al., 2011; Connor et al., 2013). The intervention had an educationally important effect for all students, boys and girls, regardless of whether they had strong or weak incoming mathematics or vocabulary skills, or attended high poverty school, keeping in mind that almost half of the children came from higher poverty homes. These findings, along with others in mathematics, reading and other content areas, provide accumulating evidence that CXI interaction effects are pervasive (Burns et al., 2010; Connor et al., 2012; Connor et al., 2013) and likely explain much of the variability in achievement observed in U.S. schools. Thus, greater attention to CXI interactions when designing and providing instruction in the early grades may have the potential to improve overall achievement for all children and form a solid foundation as students strive to become college and career ready. This is particularly important as new, comprehensive, and more rigorous learning standards in mathematics and literacy continue to be implemented and evaluated in the Common Core State Standards Initiative (Common Core State Standards Initiative, 2010).

Acknowledgements

We would like to thank the entire ISI Project team, including Barry Fishman and Christopher Schatschneider, as well as the children, parents, teachers, and school administrators without whom this research would not have been possible. We also acknowledge the contributions of Leigh McLean to an early version of this manuscript. This study was funded by "Child by Instruction Interactions: Effects of Individualizing Instruction" grant R01HD48539 and grant R21HD062834 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development; and by grants R305H04013 and R305B070074, R305A130517/R305A160404 and, in part, R305F100027 from the U.S. Department of Education, Institute of Education Sciences. The opinions expressed are ours and do not represent views of the funding agencies. ISI-Math is not currently commercially available although, as developers of the program, Dr. Crowe and Dr. Connor may have a potential conflict of interest in the future.

Appendix A. Examples of CASA items

Part One - Numeration
Number Recognition, Placement, and Patterns
Listen for directions from your teacher. (*students are told to write in numbers*)

1.	**2.**

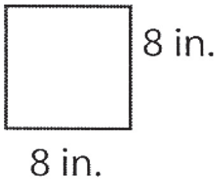
Fill in the blanks with the missing numbers

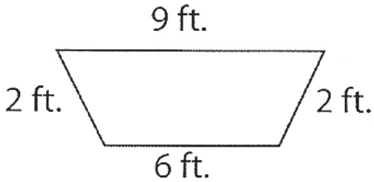
4. 14, 16, 18, __, __
5. 60, 65, 70, __, __
6. 51, 54, 57, __, __
7. 20, 30, 40, __, __

Write the number in the blank.

8. _______ two thousand six hundred
9. _______ ninety-four
10. _______ninety-six thousand, three hundred
11. _______forty-seven thousandths

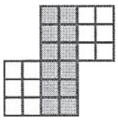
Part Two - Geometry
Area and Perimeter
Write the perimeter.

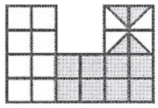




71. Perimeter = ______inches **72.** Perimeter = ________________feet

Write the area of the shaded region.

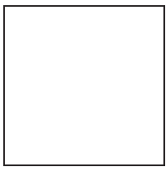




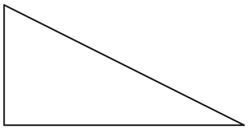
73. Area = __________ square units **74.** Area = __________ square units

Write the area.

6 in.



4 in.



3 in.

5 in. **7 in.**

75. Area = ___inches **76.** Area = ___ inches **77.** Area = ___inches

References

Agodini, R., Harris, B., Atkins-Burnett, S., Heaviside, S., Novak, T., & Murphy, R. (2009). Achievement effects of four early elementary school math curricula: Findings from first graders in 39 schools (Vol. NCEE 2009-4052). *Washington DC: National center for education evaluation and regional assistance*.

Al Otaiba, S., Connor, C. M., Folsom, J. S., Greulich, L., Meadows, J., & Li, Z. (2011). Assessment data-informed guidance to individualize kindergarten reading instruction: Findings from a cluster-randomized control field trial. *The Elementary School Journal*, *111*(4), 535–560. http://dx.doi.org/10.1086/659031.

Alcock, L., Ansari, D., Batechelor, S., Bison, M., De Smedt, B., Gilmore, C., ... Weber, K. (2016). Challenges in mathematical cognition: A collaboratively-derived research agenda. *Journal of Numerical Cognition*, *2*(1), 20–41. http://dx.doi.org/10.5964/jnc.v2i1.10.

Arnup, J., Cheree, M., John, R., & Louise, A. M. (2013). Cognitive style and gender differences in children's mathematics achievement. *Educational Studies*, *39*(3), 355–368.

Aunola, K., Leskinen, E., Onatsu-Arvilommi, T., & Nurmi, J. E. (2002). Three methods for studying developmental change: A case of reading skills and self-concept. *The British Journal of Educational Psychology*, *72*(Pt 3), 343–364.

Baker, S., Gersten, R., & Lee, D. S. (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *The Elementary School Journal*, *103*(1), 51–73.

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(5), 1860–1863. http://dx.doi.org/10.1073/pnas.0910967107.

Berch, D. B., & Mazzocco, M. M. M. (Eds.), (2007). *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities*. Baltimore, MD: Brookes Publishers.

Boaler, J. (2002). *Experiencing school mathematics: Traditional and reform approaches to teaching and their impact on student learning*. Mahwah, NJ: Lawrence Erlbaum.

Bos, C., Mather, N., Narr, R. F., & Babur, N. (1999). Interactive, collaborative professional development in early literacy instruction: Supporting the balancing act. *Learning Disabilities Research and Practice*, *14*(4), 227–238.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner, & W. Damon (Vol. Eds.), *Handbook of child psychology: Theoretical models of human development*(6th ed.). *Vol. 1*. *Handbook of child psychology: Theoretical models of human development* (pp. 793–828). Hoboken, NJ: John Wiley & Sons.

Bryant, B. R., Bryant, D. P., Kethley, C., Kim, S. A., Pool, C., & Seo, Y. J. (2008). Preventing mathematics difficulties in the primary grades: The critical features of instruction in textbooks as part of the equation. *Learning Disability Quarterly*, *31*(1), 21–35.

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

Carlisle, J. F., Kelcey, B., Rowan, B., & Phelps, G. (2011). Teachers' knowledge about early reading: Effects on students' gains in reading achievement. *Journal of Research on Educational Effectiveness*, *4*(4), 289–321.

Clements, D. H., Sarama, J., Wolfe, C. B., & Spitler, M. E. (2013). Longitudinal evaluation of a scale-up model for teaching mathematics with trajectories and technologies: Persistence of effects in the third year. *American Educational Research Journal*, *50*, 812–850. http://dx.doi.org/10.3102/0002831212469270.

Common Core State Standards Initiative (2010). *Common core state standards for English language arts & literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors.

Connolly, A. J. (2007). *KeyMath3 diagnostic assessment: Pearson*.

Connor, C. M. (2011). Child by instruction interactions: Language and literacy connections. In S. B. Neuman, & D. K. Dickinson (Eds.), *Handbook on early literacy* (pp. 256–275). (3rd ed.). New York: Guilford.

Connor, C. M. (2016). A lattice model of reading comprehension. *Child Development Perspectives*, *10*(4), 269–274. http://dx.doi.org/10.1111/cdep.12200.

Connor, C. M., Day, S. L., Phillips, B., Sparapani, N., Ingebrand, S. W., McLean, L., ... Kaschak, M. P. (2016). Reciprocal effects of self-regulation, semantic knowledge, and reading comprehension in early elementary school. *Child Development*, 1–12. http://dx.doi.org/10.1111/cdev.12570.

Connor, C. M., Morrison, F. J., Fishman, B., Giuliani, S., Luck, M., Underwood, P., ... Schatschneider, C. (2011). Testing the impact of child characteristics X instruction interactions on third graders' reading comprehension by differentiating literacy instruction. *Reading Research Quarterly*, *46*(3), 189–221. http://dx.doi.org/10.1598/RRQ.46.3.1.

Connor, C. M., Morrison, F. J., Fishman, B., Ponitz, C. C., Glasney, S., Underwood, P., ... Schatschneider, C. (2009). The ISI classroom observation system: Examining the literacy instruction provided to individual students. *Educational Researcher*, *38*(2), 85–99. http://dx.doi.org/10.3102/0013189X09332373.

Connor, C. M., Morrison, F. J., Fishman, B. J., Crowe, E. C., Al Otaiba, S., & Schatschneider, C. (2013). A longitudinal cluster-randomized controlled study on the accumulating effects of individualized literacy instruction on students' reading from first through third grade. *Psychological Science*, *24*(8), 1408–1419. http://dx.doi.org/10.1177/0956797612472204.

Connor, C. M., Morrison, F. J., Fishman, B. J., Schatschneider, C., & Underwood, P. (2007). The early years: Algorithm-guided individualized reading instruction. *Science*, *315*(5811), 464–465. http://dx.doi.org/10.1126/science.1134513.

Connor, C. M., Morrison, F. J., Schatschneider, C., Toste, J., Lundblom, E. G., Crowe, E., & Fishman, B. (2011). Effective classroom instruction: Implications of child characteristic by instruction interactions on first graders' word reading achievement. *Journal of Research on Educational Effectiveness*, *4*(3), 173–207. http://dx.doi.org/10.1080/19345747.2010.510179.

Connor, C. M., Rice, D. C., Canto, A. I., Southerland, S. A., Underwood, P., Kaya, S., ... Morrison, F. J. (2012). Child characteristics by science instruction interactions in second and third grade and their relation to students' content-area knowledge, vocabulary, and reading skill gains. *The Elementary School Journal*, *113*(1), 52–75. http://dx.doi.org/10.1086/665815.

Corbett, A. T., Koedinger, K. R., & Hadley, W. (2001). Cognitive tutors: From the research classroom to all classrooms. In P. S. Goodman (Ed.), *Technology enhanced learning: Opportunities for change* (pp. 235–263). Mahwah, NJ: Erlbaum.

Davidse, N. J., De Jong, M. T., & Bus, A. G. (2014). Explaining common variance shared by early numeracy and literacy. *Reading and Writing: An Interdisciplinary Journal*, *27*, 631–648.

Dee, A. (2010). Preservice teacher application of differentiated instruction. *The Teacher Educator*, *46*, 53.

Desoete, A., Ceulemans, A., De Weerdt, F., & Pieters, S. (2012). Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. *British Journal of Educational Psychology*, *82*(1), 64–81. http://dx.doi.org/10.1348/2044-8279.002002.

Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist*, *23*(2), 167–180.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., & Klebanov, P. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428–1446.

Edwards, C. (2006). Influences of experiences and training on effective teaching practices to meet the needs of diverse learners in schools. *Educator*, *126*, 580.

Engel, M., Claessens, A., & Finch, M. A. (2013). Teaching students what they already know? The (mis)alignment between mathematics instructional content and student knowledge in kindergarten. *Educational Evaluation and Policy Analysis*, *35*(2), 157–178. http://dx.doi.org/10.3102/0162373712461850.

Fosnot, C. T., & Dolk, M. L. A. M. (2001). *Young mathematicians at work*. Portsmouth, NH: Heinemann.

Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review*, *33*(2), 188–192.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Phillips, N. B., Karns, K., & Dutka, S. (1997). Enhancing students helping behavior during peer-mediated instruction with conceptual mathematical explanations. *The Elementary School Journal*, *97*(3), 223–249.

Fuchs, P., Seethaler, C., Fletcher, F., & Hamlett (2009). The effects of strategic counting instruction, with and without deliberate practice, on number combination skill among students with mathematics difficulties. *Learning and Individual Differences*, *20*, 89–100.

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, H. D. (2013). Adolescents' functional numeracy is predicted by their school entry number system knowledge. *PloS One*, *8*(1), e54651. http://dx.doi.org/10.1371/journal.pone.0054651.

Gravemeijer, K. P. (2014). Transforming mathematics education: The role of textbooks and teachers. *Transforming mathematics instruction* (pp. 153–172). Springer International Publishing.

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*(2), 665–668.

Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, *28*(5), 524–549.

Hiebert, J. (1999). Relationships between research and the NCTM standards. *Journal for Research in Mathematics Education*, *30*(1), 3–19.

Hiebert, J., Gallimore, R., & Stigler, J. W. (2002). A knowledge base for the teaching profession: What would it look like and how can we get one? *Educational Researcher*, *31*(5), 3–15.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, *42*(2), 371–406.

Jones, S. M., Brown, J. L., & Aber, L. (2011). Two-year impacts of a universal school-based social-emotional and literacy intervention: An experiment in translational developmental research. *Child Development*, *82*(2), 533–554. http://dx.doi.org/10.1111/j.1467-8624.2010.01560.x.

Jordan, N. C., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Differences*, *20*(2), 82–88.

LeFevre, J. A., Fast, L., Skwarchuk, S. L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., & Penner-Wilger, M. (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child Development*, *81*(6), 1753–1767.

Leung, K. C. (2015). Preliminary empirical model of crucial determinants of best practice for peer tutoring on academic achievement. *Journal of Educational Psychology*, *107*(2), 558–579.

Li, X., Chi, L., DeBey, M., & Baroody, A. J. (2015). A study of early childhood mathematics teaching in the United States and China. *Early Education and Development*, *26*(3), 450–478.

Lubienski, S. T. (2004). Traditional or standards-based mathematics? The choices of students and parents in one district. *Journal of Curriculum and Supervision*, *19*(4), 338–365.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mazzocco, M.M.M., Crowe, E.C., Calhoon, M.B., & Connor, C.M., (in preparation). Child-characteristic X instruction interactions in mathematics: The effect of peer-

assisted learning strategies.
Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011a). Impaired acuity of the approximate number system underlies mathematical learning disability. *Child Development*, *82*(4), 1224–1237.
Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011b). Preschoolers' precision of the approximate number system predicts later school mathematics perfomance. *PloS One*, *6*(9), e23749. http://dx.doi.org/10.1371/journal.pone.0023749.
Mazzocco, M. M. M., Murphy, M. M., Brown, E. C., Rinne, L., & Herold, K. H. (2013a). Persistent consequences of atypical early number concepts. *Frontiers in Psychology*, *4*. http://dx.doi.org/10.3389/fpsyg.2013.00486.
Mazzocco, M. M. M., & Myers, G. F. (2003). Complexities in identifying and defining mathematics learning disability in the primary school-age years. *Annals of Dyslexia*, *53*, 218–253.
Mazzocco, M. M. M., Myers, G. F., Lewis, K. E., Hanich, L. B., & Murphy, M. M. (2013b). Limited knowledge of fraction representations differentiates middle school students with mathematics learning disability (dyscalculia) versus low mathematics achievement. *Journal of Experimental Child Psychology*, *115*(2), 371–387. http://dx.doi.org/10.1016/j.jecp.2013.01.005.
Mazzocco, M. M. M., & Räsänen, P. (2013). Contributions of longitudinal studies to evolving definitions and knowledge of developmental dyscalculia. *Trends in Neuroscience and Education*, *2*(2), 65–73. http://dx.doi.org/10.1016/j.tine.2013.05.001.
Mazzocco, M. M. M., & Thompson, R. E. (2005). Kindergarten predictors of math learning disability. *Learning Disabilities Research and Practice*, *20*, 142–156.
Morgan, P. L., Farkas, G., & Maczuga, S. (2015). Which instructional practices most help first-grade students with and without mathematics difficulties? *Educational Evaluation and Policy Analysis*, *37*(2), 184–205. http://dx.doi.org/10.3102/0162373714536608.
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.
Murphy, M. M., Mazzocco, M. M. M., Hanich, L., & Early, M. C. (2007). Cognitive characteristics of children with mathematics learning disability (MLD) vary as a function of the cut-off criterion used to define MLD. *Journal of Learning Disabilities*, *40*, 458–478.
NAEP (2015). *The nation's report card: Reading highlights*. Washington DC: National Center for Education Statistics.
National Council of Teachers of Mathematics (NCTM) (Producer) (2013, March). Principles and standards for school mathematics. Retrieved from http://www.nctm.org/standards/content.aspx?id=16909.
National Mathematics Advisory Panel (2008). Foundations for success: The final report of the National Mathematics Advisory Panel. Retrieved from http://www.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf.
Nunes, T., Bryant, P., Evans, D., & Barros, R. (2015). Assessing quantitative reasoning in young children. *Mathemtical Thinking and Learning*, *17*(2–3), 178–196.
Pearson, B. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology*, *8*, 317–344.
Petrill, S., Logan, J., Hart, S., Vincent, P., Thompson, L., Kovas, Y., & Plomin, R. (2012). Math fluency is etiologically distict from untimed math performance, decoding fluency, and untimed reading performance: Evidence from a twin study. *Journal of Learning Disabilities*, *45*(4), 371–381.
Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, *48*(1), 85–112. http://dx.doi.org/10.1016/j.jsp.2009.09.002.
Price, G. R., Mazzocco, M. M. M., & Ansari, D. (2013). Why mental arithmatic counts: Brain activation durin single digit arithmetic predicts high-school math scores. *Journal of Neuroscience*, *33*(1), 156–163.
Purpura, D. J., Baroody, A. J., & Lonigan, C. J. (2013). The transition from informal to formal mathematical knowledge: Mediation by numeral knowledge. *Journal of Educational Psychology*, *105*(2), 453–464. http://dx.doi.org/10.1037/a0031753.
Purpura, D. J., Hume, L. E., Sims, D. M., & Lonigan, C. J. (2011). Early literacy and early numeracy: The value of including early literacy skills in the prediction of numeracy development. *Journal of Experimental Child Psychology*, *110*, 647–658. http://dx.doi.org/10.1016/j.jecp.2011.07.004.
Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R., & du Toit, M. (2004). *HLM6: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.
Resendez, M., & Azin, M. (2005). *The relationship between using Saxon elementary and middle school math and student performance on Georgia statewide assessments*. Jackson, WY: Pres Associates.
Resendez, M., & Sridiharan, S. (2006). *The relationship between using Saxon elementary school math and student performance on Texas statewide assessments*. Jackson, WY: PRES Associates.
Reys, R. E. (2001). Curricular controversy in the math wars: A battle without winners. *Phi Delta Kappan*, *83*(3), 255–258.
Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., Nathanson, L., & Brock, L. L. (2009). The contribution of children's self-regulation and classroom quality to children's adaptive behaviors in the kindergarten classroom. *Developmental Pscyhology*, *45*(4), 958–972.
Riordan, J. E., & Noyce, P. E. (2001). The impact of two standards-based mathematics curricula on student achievement in Massachusetts. *Journal for Research in Mathematics Education*, *32*(4), 368–398.
Ritchie, W. C., & Bhatia, T. K. (Eds.), (1999). *Handbook of child language acquisition*. San Diego: Academic Press.
Rivera, F. D. (2014). From math drawings to algorithms: Emergence of whole number operations in children. *ZDM Mathematics Education*, *46*(1), 59–77.
Rivera-Batiz, L. (1992). *Essays on trade, technology, and economic growth. (Ph D), University of Chicago*.
Robinson-Cimpian, J. P., Lubienski, S. T., Ganley, C. M., & Copur-Gencturk, Y. (2014). Teachers' perceptions of students' mathematics proficiency may exacerbate early gender gaps in achievement. *Developmental Psychology*, *50*(4), 1262–1281.
Rogers, K. B. (2000). Lessons learned about educating the gifted and talented a synthesis of the research on educational practice. *The Gifted Child Quarterly*, *51*(4), 382–396.
Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. *Handbook of research on mathematics teaching and learning* (pp. 334–370). .
Schoenfeld, A. H. (2004). The math wars. *Educational Policy*, *18*(1), 253–286.
Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., ... Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, *23*(7), 691–697. http://dx.doi.org/10.1177/0956797612440101.
Sood, S., & Jitendra, A. K. (2007). A comparative analysis of number sense instruction in reform-based and traditional mathematics textbooks. *The Journal of Special Education*, *41*(3), 145–157.
Stallings, J. (1980). Allocated academic learning time revisited, or beyond time on task. *Educational Researcher*, *9*(11), 11–16.
Stoet, G., & Geary, D. C. (2015). Sex differences in academic achievement are not related to political, economic, or social equality. *Intelligence*, *48*, 137–151.
Turner, E. E., Drake, C., McDuffie, A. R., Aguirre, J., Bartell, T. G., & Foote, M. Q. (2012). Promoting equity in mathematics teacher preparation: A framework for advancing teacher learning of children's multiple mathematics knowledge bases. *Journal of Mathematics Teacher Education*, *15*(1), 67–68.
Van den Heuvel-Panhuizen, M. (2010). Reform under attack – Forty years of working on better mathematics education thrown on the scrapheap? No way. In L. Sparrow, B. Kissane, & C. Hurst (Eds.), *Shaping the future of mathematics education: Proceedings of the 33rd annual conference of the Mathematics Education Research Group of Australasia* (pp. 1–25). Fremantle: MERGA.
VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197–221. http://dx.doi.org/10.1080/00461520.2011.611369.
Willcutt, E. G., & Pennington, B. F. (2000). Comorbidity of reading disability and attention-deficit/hyperactivity disorder: Differences by. *Journal of Learning Disabilities*, *33*(2), 179.
Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson-III tests of achievement*. Itasca, IL: Riverside.
Yoshikawa, H., & Hsueh, J. (2001). Child development and public policy: Toward a dynamic systems perspective. *Child Development*, *72*(6), 1887–1903. http://dx.doi.org/10.1111/1467-8624.00384.