Improving Children's Understanding of Mathematical Equivalence: An Efficacy Study

Jodi L. Davenport*[1], Yvonne S. Kao[1], Kristen N. Johannes[1], Caroline Byrd Hornburg[2], and Nicole M. McNeil[3]

[1]WestEd

[2]Virginia Tech

[3]University of Notre Dame

Author Note

*Corresponding author is Jodi L. Davenport: jdavenp@wested.org

This document is a pre-proof draft. Davenport, J. L., Kao, Y., Johannes, K., Hornburg, C. B., & McNeil, N. M. (2022). Improving children's understanding of mathematical equivalence: An efficacy study. Journal of Research on Educational Effectiveness. DOI: https://doi.org/10.1080/19345747.2022.2144787.

Available at: https://www.tandfonline.com/doi/abs/10.1080/19345747.2022.2144787

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A150088 to WestEd. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Abstract

A vast majority of elementary students struggle with the core, pre-algebraic concept of mathematical equivalence. The Improving Children's Understanding of Equivalence (ICUE) intervention integrates four research-based strategies to improve outcomes for second grade students: (1) introducing the equal sign before arithmetic, (2) non-traditional arithmetic practice, (3) concreteness fading exercises, and (4) comparison and explanation. In a large-scale randomized control trial in California public schools, 132 second grade teachers were randomly assigned to either use the ICUE intervention or an active control consisting of non-traditional arithmetic practice alone. Using data from 121 teachers in the analytic sample, the study found that students in the intervention group outperformed students in the active control on proximal and transfer measures of equivalence with no observable trade-offs in computational fluency. The findings suggest that the ICUE intervention helps students construct a robust understanding of mathematical equivalence, a critical precursor to success in algebra.

Improving Children's Understanding of Mathematical Equivalence: An Efficacy Study

Introduction

Understanding mathematical equivalence is a critical precursor to advanced mathematical concepts and algebraic thinking. However, the majority of elementary students struggle to grasp equivalence, and traditional instruction may perpetuate unhelpful conceptions. To help students construct a formal understanding of equivalence, McNeil and her team (Hornburg, Brletic-Shipley, Matthews, & McNeil, 2021; McNeil, Hornburg, Brletic-Shipley, & Matthews, 2019a) developed a research-based intervention for second grade students, *Improving Children's Understanding of Equivalence (ICUE)*. The current efficacy study evaluated whether the sizable effects found in smaller studies would be similarly found in a large-scale randomized control trial conducted in public school classrooms across the state of California.

Background

Importance of instruction targeting mathematical equivalence

Algebra is widely viewed as a gatekeeper to college and career success (Adelman, 2006; Ma, 2001; National Mathematics Advisory Panel, 2008). Unfortunately, by eighth grade, only 34% of students in the United States have reached proficiency in mathematics, and many students are still building the conceptual foundations necessary for algebraic reasoning (Hussar et al., 2020; Loveless, 2008). In response, researchers, educators, and policymakers advocate for instruction to enhance prerequisite understandings in early elementary school that prepare students for later success in algebra (Blanton & Kaput, 2005; Blanton et al., 2015, 2019; Jacobs, Franke, Carpenter, Levi, & Battey, 2007; National Mathematics Advisory Panel, 2008).

Mathematical equivalence, the relationship between two interchangeable quantities, is fundamental to algebraic reasoning (e.g., Blanton et al., 2015; Charles, 2005; Fyfe, Matthews, Amsel, McEldoon, & McNeil, 2018; Jones, Inglis, Gilmore, & Dowens, 2012; McNeil, 2014). Understanding mathematical equivalence involves knowing that numbers, measurements, and expressions can be represented in a variety of equivalent ways and that the equal sign reflects the equivalence relation "is equal to" (Jacobs et al., 2007; Knuth, Stephens, McNeil, & Alibali, 2006). Formal understanding of mathematical equivalence in second grade predicts future mathematical competence, such as achievement on a standardized mathematics assessment in third grade (McNeil, Hornburg, Devlin, Carrazza, & McKeever, 2019b), proficiency with algebraic procedures in fourth grade (Matthews & Fuchs, 2020), and algebra readiness in sixth grade (Hornburg, Devlin, & McNeil, 2022). Understanding equivalence further predicts performance solving algebraic equations in middle school (Alibali, Knuth, Hattikudur, McNeil, & Stephens, 2007; Knuth et al., 2006) and college (Fyfe, Matthews, & Amsel, 2020).

Despite its importance as a foundation of mathematical thinking, a majority of US students do not reason with or apply concepts of mathematical equivalence when presented with symbolic mathematics problems (e.g., Blanton et al., 2019; Byrd, McNeil, Chesney, & Matthews, 2015; Fyfe et al., 2018). Compared with 90% of Chinese students, fewer than 20% of US students (ages 7-11) successfully solve a common test of mathematical equivalence reasoning that involves equations with operations on both sides of the equal sign, e.g., $3 + 7 + 5 = 3 + \_$, (Li, Ding, Capraro, & Capraro, 2008; Perry, Church, & Goldin-Meadow, 1988). US students not only solve the problems incorrectly, but also encode equations that have operations on both sides of the equal sign, e.g., $3 + 7 + 5 = 3 + \_$, as if they were written with operations only on the left side, e.g., $3 + 7 + 5 + 3 = \_$, (Alibali, Phillips, & Fischer, 2009; McNeil & Alibali, 2004; Rittle-Johnson & Alibali, 1999), and define the equal sign operationally as “calculate the total” or “put the answer” rather than a relational symbol that indicates mathematical equivalence (e.g., Baroody & Ginsburg, 1983; McNeil & Alibali, 2005a; Powell & Fuchs, 2010).

Why is mathematical equivalence hard to teach?

The change-resistance account of students' conceptions of mathematical equivalence suggests that difficulties in building a deep, formal understanding of mathematical equivalence stem from logical, but inappropriate, generalization of knowledge constructed from narrow experiences in early mathematics (McNeil, 2014; McNeil & Alibali, 2005b; see also McNeil, Hornburg, Fuhs, & Connor, 2017). Repeated practice solving arithmetic problems in traditional formats, coupled with a lack of explicit instruction on equivalence, leads students to develop long-term memory representations of equations that focus on operations (e.g., the calculations involved) rather than relations (specifically, that the equal sign expresses a relationship between the two sides of an equation). In early elementary school, "typical" arithmetic problems, e.g., $2 + 3 = 5$, present operations on the left of the equal sign and the "answer" on the right (McNeil et al., 2006; Powell, 2012). Children detect and extract patterns from these examples, construct long-term memory representations, and incorrectly generalize "operational patterns" (e.g., the equal sign means "find the answer") to later mathematics practice (cf. Jacobs et al., 2007). Though operational patterns might speed computation in "typical" arithmetic problem solving, as patterns become entrenched, students may mistakenly transfer their knowledge to similar, but not applicable, problem types (e.g., Bruner, 1957; Chase & Simon, 1973; Mickey & McClelland, 2014). McNeil et al. (2019b) used a longitudinal design to study the potential causes and consequences of early understanding of mathematical equivalence and found results consistent with predictions of the change-resistance account; students who relied on the operational patterns needed more support for learning mathematical equivalence than those who were incorrect for other reasons. Similarly, Byrd and colleagues (2015) found that students who held an arithmetic-specific view of the equal sign (e.g., "add all the numbers") learned less from lessons on mathematical equivalence than students with other incorrect views not explicitly tied to operations (e.g., "the end of the problem").

As entrenched patterns cannot be remedied by a quick fix (Seidenberg & Zevin, 2006), many prior instructional attempts to help children construct a formal understanding of mathematical equivalence have had mixed results (e.g., Jacobs et al., 2007; McNeil, Fyfe, Petersen, Dunwiddie, & Brletic-Shipley, 2011). Some interventions increase performance on proximal measures, but not on transfer problems that differ on surface features (e.g., Alibali, 1999; Perry, 1991). Others temporarily increase learning and transfer, but result in a re-emergence of incorrect ways of thinking a few weeks after initial learning (e.g., Cook, Mitchell, & Goldin-Meadow, 2008; McNeil & Alibali, 2000). Finally, some interventions successfully improve student understanding, but require either math education researchers going into the schools to teach lessons throughout the school year (e.g., Blanton et al., 2019), or the combination of multiple professional development sessions, highly motivated teachers, and an unsustainable amount of class time devoted to the intervention (e.g., Carpenter & Levi, 2000; Schliemann, 2007).

Design features of the Improving Children's Understanding of Equivalence (ICUE) intervention

The ICUE intervention was designed for second grade because the change-resistance account suggests that early intervention is important to prevent children from developing and relying on operational patterns. Openness to understanding mathematical equivalence appears to decrease between the ages of seven and nine (McNeil, 2007), and earlier understanding of mathematical equivalence in elementary school is associated with better algebra readiness in middle school, even after controlling for important individual differences (Hornburg et al., 2022).

Non-traditional arithmetic practice. A key component of the ICUE intervention is non-traditional arithmetic practice. The change-resistance account assumes that narrow exposure to "traditional" arithmetic problems leads students to develop operational patterns that impede formal understanding of mathematical equivalence. Presenting students with a wider variety of arithmetic problems appears to prevent the formation and/or activation of these operational patterns and leads to improved understanding of mathematical equivalence. Examples of this type of arithmetic practice include (a) presenting arithmetic problems in non-traditional problem formats that put operations on the right side of the equal sign, e.g., $\_ = 9 + 8$, (McNeil et al., 2011); (b) organizing problems into practice sets based on equivalent values, e.g., $2 + 5 = \_$, $3 + 4 = \_$, $6 + 1 = \_$, (McNeil et al., 2012); and (c) using relational phrases, such as "is equal to" and "is the same amount as," in place of the equal sign in practice problems (Chesney, McNeil, Petersen, & Dunwiddie, 2018). Students who practice arithmetic in these non-traditional ways construct a better understanding of mathematical equivalence (McNeil, Fyfe, & Dunwiddie, 2015). Moreover, the benefits of non-traditional arithmetic practice are robust, persisting five to six months after the intervention, and there have not been any observed trade offs with computational fluency. Consistent with the change-resistance account's theory of change, children who receive non-traditional arithmetic practice are less likely to rely on operational patterns when solving mathematical equivalence problems (e.g., adding up all the numbers), less likely to encode mathematical equivalence problems as if they were typical addition problems, and less likely to write "operations = answer" problem formats when generating mathematics problems on their own. A mediation analysis conducted by McNeil et al. (2015) showed that random assignment to non-traditional arithmetic practice improves understanding of mathematical equivalence by decreasing children's reliance on the operational patterns.

Though exposure to non-traditional arithmetic practice improves student performance, non-traditional practice alone is not sufficient for cultivating a deep and lasting understanding of mathematical equivalence for the majority of students. The full

ICUE intervention was developed to address this gap and includes three additional, research-based components: (1) exposure to the equal sign in non-arithmetic contexts (e.g., 28 = 28) before introducing arithmetic expressions within equations, (2) concreteness-fading exercises that first present concrete, real-world, relational contexts (e.g., sharing stickers, balancing a scale) before fading into the corresponding abstract mathematical symbols (e.g., numerals, operators, the equal sign), and (3) activities that require students to compare and explain different problem formats and problem-solving strategies. Below we describe how each additional component supports learning.

Introduction of equal sign before arithmetic. Children informally interpret addition as a unidirectional process before the start of formal schooling (Baroody & Ginsburg, 1983) and apply the operational patterns to arithmetic problems as early as first grade (e.g., Falkner, Levi, & Carpenter, 1999). To promote correct, relational understandings, the first ICUE lessons introduce the equal sign and concept of equivalence outside of arithmetic (e.g., using the equal sign to compare the number of dogs and bones, lightbulbs and lamps, or two boxes of stickers) before presenting a variety of non-traditional arithmetic problem formats. McNeil (2008) showed that children learn more from instruction on the equal sign when the equal sign is presented outside of (versus within) the context of arithmetic, and children in that study who were the most reliant on operational patterns reaped the largest benefits from seeing the equal sign outside of an arithmetic context (McNeil, 2008). Other studies have also provided indirect evidence that children benefit from constructing a relational definition of the equal sign prior to exposure to symbolic arithmetic problems (Baroody & Ginsburg, 1983; Denmark, Barco, & Voran, 1976; Li et al., 2008; Renwick, 1932).

Concreteness fading exercises. As elementary school children demonstrate a correct understanding of mathematical equivalence in realistic contexts, many experts recommend beginning with concrete representations of mathematics concepts and slowly fading to more abstract representations (Bruner, 1966; Fyfe, McNeil, Son, & Goldstone, 2014;

Goldstone & Son, 2005; Gravemeijer, 2002). These “concreteness fading” exercises are the third component of ICUE. Outside of formal, symbolic arithmetic, students have less difficulty defining the word “equal” correctly, identifying addend pairs that are “equal to” one another (e.g., Rittle-Johnson & Alibali, 1999) or solving mathematical equivalence problems with concrete materials (e.g., wooden blocks), instead of with numbers written in symbolic form (Sherman & Bisanz, 2009). Fyfe, McNeil, and Borjas (2015) found that children in a concrete-to-abstract fading condition constructed a better understanding of mathematical equivalence than children in concrete only, abstract only, or abstract-to-concrete progression conditions. Even when incorrect, children in the concrete-to-abstract fading condition were least reliant on operational patterns after instruction.

Comparison and explanation. The final component of ICUE requires students to compare and explain different problem formats and problem-solving strategies. Explicit comparison requires students to articulate and justify generalizations about the underlying structure and properties of arithmetic that go beyond operational patterns (Carpenter, Franke, & Levi, 2003, see also Blanton et al., 2015; Falkner et al., 1999; Hattikudur & Alibali, 2010; Jacobs et al., 2007). Explanation of why correct strategies are correct and why incorrect strategies are incorrect facilitates children’s understanding of mathematical equivalence (e.g., Jacobs et al., 2007; Rittle-Johnson, 2006; Siegler, 2002). Children asked to compare and explain tend to invent new strategies rather than sticking rigidly to the erroneous strategy of adding all the numbers (Rittle-Johnson, 2006; Siegler, 2002).

The Improving Children’s Understanding of Equivalence (ICUE) intervention

The ICUE intervention combines the four components into 32 supplementary lessons and paper-based activities. Students participate in two 15-20 minute sessions per week over 16 weeks in their regular classrooms, for 8-10 total hours of instructional time over the school year. All lessons are used in addition to regular math instruction. ICUE materials include teacher manuals, individual student workbooks, manipulatives (e.g., stickers and balance scales), and assessment items. The teacher manual includes a summary of students' difficulties understanding mathematical equivalence, the proposed scope, sequence, and schedule of the activities, a list of required materials, scripted lessons for each activity, instructions for assessments, and tips for successful implementation. Students receive their own workbooks to use during the lessons and practice activities. Figure 1 shows pictures of the materials and Online Appendix A provides examples from the teacher manual of the sequence, lessons, and practice activities.

Building on prior results

A small cluster randomized study demonstrated positive outcomes for students receiving the ICUE intervention. Eight second grade classrooms in under-resourced Catholic schools were randomly assigned to use either the full ICUE intervention or an active control of nontraditional arithmetic practice activities. The active control condition was the practice workbook created by McNeil, Fyfe, and Dunwiddie (2015), that included non-traditional arithmetic problems with operators on the right side, problems where the equal sign was replaced with the words "is equal to," and some problem sets that were organized by equivalent sums. Analyses at the classroom level found that full ICUE classrooms performed better than active control classrooms on both proximal measures of understanding of mathematical equivalence ($\eta^2 = 0.70$) and transfer problems ($\eta^2 = 0.67$), (McNeil et al., 2019a). All eight teachers in the study were first- and second-year teachers who had Bachelor's degrees in non-teaching fields and were completing their Master of Education through the ACE Teaching Fellows program, a post-baccalaureate "service through teaching" program.

The current study sought to replicate and extend these promising findings. We conducted a large-scale randomized control trial in public schools with diverse student populations and in classrooms of teachers with differing backgrounds. In addition, we extend the findings by including more distal, independently-developed measures of mathematical equivalence. In our replication, we again compared the ICUE intervention to the "non-traditional arithmetic" active control condition used by McNeil, Hornburg, Brletic-Shipley, and Matthews (2019a) for pragmatic and ethical reasons. First, as the ICUE intervention is designed to supplement, not replace, business-as-usual math instruction, we wanted to ensure any positive effects were not simply due to more time spent on math practice. Second, as the active control has been shown to improve student learning over business as usual in a study meeting What Works Clearinghouse recommendations without reservations (McNeil, 2014), we felt it would be inappropriate to deny the control condition a supplemental intervention known to have positive effects. We wanted to offer all teachers in the study an opportunity to use a supplementary intervention with evidence of helping children construct relational understanding of mathematical equivalence. Finally, teachers who have sufficient interest in and conceptual understanding of the Common Core Mathematics Standards may already incorporate non-traditional arithmetic into their business-as-usual practices.

Similar to the ICUE intervention, the active control materials included 32 lessons and paper-based activities to be used in two 15-20 minute sessions per week over 16 weeks. All lessons were to be used in addition to regular math instruction. The active control materials included teacher manuals, individual student workbooks, and assessment items, but did not include the additional manipulatives such as balance scales. Active control lessons included non-traditional arithmetic practice, but did not include the other three ICUE design features, e.g., introduction of the equal sign before arithmetic, concreteness fading exercises, or comparison and explanation. See the Online Appendix B for examples of the sequence of activities and examples of lessons as reflected in the teacher manual.

Research Questions

Our study was designed to address the following two research questions based on past work: 1) Does the ICUE intervention lead to improved performance compared to an active control (i.e., non-traditional arithmetic practice only) on measures of mathematical equivalence controlling for student demographics (e.g., gender, underrepresented student status, English learner) and teacher background? 2) Do improvements in performance on mathematical equivalence come at the expense of computational fluency?

Given the potential theoretical and practical benefits of considering heterogeneity of intervention effects (Bryan, Tipton, & Yeager, 2021), we also sought to address a third, more exploratory question: 3) Do treatment effects interact with participant characteristics? Although we did not have specific predictions about specific participant characteristics that might attenuate or boost ICUE intervention effects, testing for moderators of intervention effects can be a powerful tool for both identifying potential causal mechanisms, and providing nuanced guidance to educators seeking to implement the intervention (Bryan et al., 2021).

Method

Study design

This study was a cluster-randomized trial of the ICUE intervention compared to an active control condition that had already been shown to improve students' understanding of mathematical equivalence relative to traditional arithmetic practice. Our randomized controlled trial (RCT) study was designed to meet the What Works Clearinghouse (WWC) Group Design Standards, Version 4.1. The goal of ICUE is to be an "out of the box" supplementary intervention that can be used to supplement learning in any mathematics classroom regardless of the business-as-usual instruction. We used a random process that assigned teachers to conditions entirely by chance and every teacher had an equal chance of being assigned to the treatment versus control condition.

Participants

The study took place in second grade classrooms in public school districts across California. A total of 132 second grade teachers and 3006 second grade students from 53 schools in 7 districts participated. Random assignment was conducted at the teacher level after the start of the school year when class rosters were set. Each teacher had a 50% probability of being assigned to either condition. Sixty-nine teachers (1567 students) were assigned to the treatment condition and 63 teachers (1439 students) were assigned to the control condition.

The analytic sample contains data from 121 teachers (60 treatment, 61 control) representing 53 schools in 7 districts. This represents 13.0% cluster attrition from the treatment condition and 3.2% from the control condition, for a total cluster attrition of 8.3%. According to the What Works Clearinghouse Evidence Review Protocol for Elementary School Mathematics Interventions Version 2.0 (What Works Clearinghouse, Institute of Education Sciences, 2012), the level of attrition from the study is unlikely to be a source of bias. Per our Institutional Review Board protocol, teachers were able to cease participation at any time without explanation. However, anecdotally, most teachers left due to reassignment or leaves of absence.

Four teachers identified as male and 117 identified as female. Overall, this was an experienced sample of teachers, with 7.4% reporting less than three years of mathematics teaching experience, 15.7% reporting 3-8 years of mathematics teaching experience, and 76.9% reporting more than nine years of mathematics teaching experience. All teachers in the sample reported having at least a Bachelor's degree and 59.5% reported having at least a Master's degree. Most teachers (74.4%) had studied elementary or early childhood education for their degree(s); no teachers had studied mathematics or mathematics education.

All classes in the analytic sample were self-contained, meaning that each second grade teacher taught all core subjects to his or her students. Eighty-two percent of teachers reported teaching between four and seven hours of mathematics per week. Teachers most commonly reported teaching five hours of mathematics instruction per week. As intended, teachers reported using ICUE to supplement, not replace, business as usual instruction.

All 2259 students (1131 treatment, 1128 control) with complete data from non-attriting classes are included in the analytic sample. This represents 18.3% attrition from non-attriting treatment clusters and 19.4% from non-attriting control clusters, for a total sub-cluster attrition of 18.9%. The demographics of students in the analytic sample are reasonably representative of public school students in California (California Department of Education, 2020a, 2020b), shown in Table 1. We did not collect data on socioeconomic status for individual students, but on average 55% (SD = 35%) of students enrolled in analytic sample schools were eligible for either free or reduced-price lunch compared to approximately 61% of students in California (California Department of Education, n.d.).

Measures

Implementation Measures

Teacher logs. Each week, teachers in both conditions completed an implementation log reporting how they used the study materials with their classes. Teachers reported if and when they taught each session of their assigned materials, how long they spent preparing for each session (in 5-minute increments), and how much class time was spent on each session (in 5-minute increments). Teachers also had the opportunity to report any additional comments or issues related to study participation.

Workbook completion. Both ICUE and the active control are supplementary interventions that rely on student workbooks as the primary mode of practice. Every lesson in both treatment and active control conditions centers student work in paper-based workbooks, so workbook completion is a strong measure of implementation fidelity. We used stratified random sampling to select four students from each class, performing a quartile split on students' pretest scores and randomly selecting one student from each quartile, and requested that teachers mail us these students' workbooks. To minimize any possibility that teachers would give more attention to selected students, teachers were not informed which students' workbooks would be requested until after all sessions were completed.

Teacher Mathematics Knowledge for Teaching measure

As the majority of elementary school teachers in California hold a Multiple Subjects credential (to teach all subjects) rather than a Single Subject credential (e.g., to specifically teach mathematics), there may be variance in knowledge of math teaching. Teachers completed the 2008 elementary school Number Concepts and Operations test from the Mathematics Knowledge for Teaching assessments developed by the Learning Mathematics for Teaching project at the University of Michigan (Hill, Ball, & Schilling, 2008). The measure provides an indicator of teachers' specialized content knowledge, ability to analyze student thinking, identify students' mathematical understandings, and represent mathematical ideas in differing ways.

Researcher-developed student outcome measures

Students completed the same paper-and-pencil assessments of understanding of mathematical equivalence used in the smaller-scale ICUE study (McNeil et al., 2019a). At posttest, students completed an assessment that included equation encoding items (4), open-ended questions asking students to name and define the equal sign (2), equation solving items (8), equation-solving transfer items (4), and addition fluency items (81). At pretest students completed the same items, with the exception of the four equation-solving transfer items. We describe these items below.

Equation encoding (pre- and post-test). This assessment includes four encoding items. To measure encoding, students are asked to reconstruct math equivalence problems after viewing each problem projected in front of the class for five seconds. A reconstruction is coded as correct if the equation appears exactly as it was shown. Cohen's Kappa, a measure of inter-rater reliability, was 0.94 at pre-test and 0.96 at post-test. Cronbach's alpha was 0.59 at pre-test and 0.65 at post-test, based on data from our analytic sample.

Equal sign naming and defining (pre- and post-test). To measure equal sign understanding, students respond to questions about the equal sign. An arrow points to an equal sign presented at the top of the page and the text reads: (1) What is the name of this math symbol? and (2) What does this math symbol mean? Teachers read each question aloud and children write their responses after each one. Students' definitions are coded for whether or not children correctly define the equal sign as a symbol of math equivalence (e.g., two amounts are the same, something is equivalent to another thing). Although many students in this age range have poor spelling, coders in the past have not had trouble determining what a given child has written, even when words are misspelled (e.g., "the toltal," "write the anser next," the same azz"). Cohen's Kappa was 0.97 at pre-test and 0.98 at post-test, and for defining the equal sign was 0.86 at pre-test and 0.84 at post-test. Because Cronbach's alpha is affected by the length of the measure, and the equal sign naming and defining test only contained two items, we do not report alpha here.

Equation-solving (pre- and post-test). This assessment contains eight equations that have operations on both sides of the equal sign (e.g., $3 + 5 + 6 = 3 + \_$). For the analysis presented in this paper, students' responses were simply scored as correct or incorrect for a maximum score of eight. Cronbach's alpha for the equation-solving assessment was 0.94, based on pre-test data from our analytic sample.

Equation-solving transfer measure (post-test only). This assessment includes four more advanced equations for students to solve. These equations included subtraction as an operation on the right side of the equal sign (e.g., $2 + 5 + 3 = 14 - \_$) as well as larger numbers (e.g., $13 + 18 = \_ + 19$) (McNeil et al., 2019). Students' responses were simply scored as correct or incorrect for a maximum score of four. The transfer measure was only given as a post-test. Cronbach's alpha for the combined equation-solving measures (maximum score of 12) was 0.94, based on post-test data from our analytic sample.

Addition fluency (pre- and post-test). This assessment includes 81 addition problems, consisting of the pairwise combinations of numbers 1 to 9. Students are given one minute to solve as many problems as possible and their final score is the total number of problems solved correctly. We used our analytic sample data to compute a split-half reliability coefficient with the Spearman-Brown correction, comparing students' performance on odd-numbered vs. even-numbered items. The reliability coefficient was 0.98 on pre-test and 0.99 on post-test.

Independently-developed student outcome measures (posttest only)

To test whether students' learning transferred to measures that were not developed by researchers involved in ICUE's development, we gave students performance tasks from the Silicon Valley Mathematics Initiative's Mathematics Assessment Collaborative (MAC). MAC partners with the Mathematics Assessment Resource Service (MARS) to develop tasks that assess core mathematical ideas and practices taught in each grade level (Darling-Hammond & Falk, 2013). Tasks require students to solve complex mathematics problems and provide open-ended explanations of their reasoning. We selected two MARS tasks that tested students' understanding of mathematical equivalence, described below. For each task, MARS provides scoring rubrics and scorer training documents, student performance statistics, and examples of common student errors. Students' responses on both tasks were hand-scored by experienced scorers that followed the standardized training, calibration, and quality assurance procedures established by MARS (Foster & Noyce, 2004).

Incredible Equations. In this task, students were asked to fill in the missing parts of equations such as $\_ + 8 + \_ = 16$ and $11 + 5 = \_ + 8$. Students were then asked to explain how they know their answer is correct. MARS administered this task to 6,305 second graders in 2007. The mean score was 6.08 out of 10 with a standard deviation of 2.5 (Mathematics Assessment Resource Service, 2007).

Agree or Disagree? In this task, students were asked if they agree or disagree with statements of equivalence: $8 + 5 = 5 + 8$ and $6 - 4 = 4 - 6$. Students were then asked to explain their answers using words, numbers, or pictures. MARS administered this task to 4,585 second graders in 2004. The mean score was 3.10 out of 6 with a standard deviation of 1.94 (Mathematics Assessment Resource Service, 2004).

MARS does not report reliability statistics on its tasks. Based on responses from our analytic sample, Cronbach's alpha is 0.76 for *Incredible Equations*, 0.6 for *Agree or Disagree?*, and 0.78 for both items combined. For this analysis, we combine both MARS tasks into a single measure.

Finally, as another check to ensure there were no trade-offs with computational fluency, students also completed the computation subtest of Level 8 of the *Iowa Test of Basic Skills (ITBS)*, a standardized measure of general mathematical reasoning. The computation subtest is designed to be completed in 20 minutes and consists of 31 multiple-choice questions that assess students' foundational understanding of quantity and the number system, operations and algebra, measurement, and geometry. Developers report a Kuder-Richardson Formula 20 median reliability coefficient of 0.84 for the Grade 2 test.

Study procedure

The procedure for ICUE Treatment and Active Control conditions were identical, differing only in the content of the materials used by teachers and students. Each teacher participated in a one-hour, online training for their assigned condition. The training reviewed the study purpose, features of the activities, and strategies for integrating the activities into their typical mathematics curriculum. During this training, researchers stressed the importance of using only the materials provided and not to discuss materials with other teachers that may be participating at their school. Teachers were also reminded that the activities should be used to supplement whatever math curriculum they were already using. That is, teachers should teach their regular math lessons as they normally would and add the ICUE or active control materials twice a week.

After the training and prior to starting the study, participating teachers completed a background survey and the online Mathematics Knowledge for Teaching measure. The intervention centers on workbook practice and teachers only received the materials for their assigned condition, e.g., workbooks for all students and balance scales for treatment classrooms. This characteristic lowers the potential for teachers in different conditions within the same school to influence one another. Teachers administered the paper-and-pencil pretest assessment described above that included equation solving items, equation encoding items, open-ended questions asking students to identify and define the equal sign, and addition fluency items. Teachers were asked to start Session 1 within one week of administering the pretest.

Teachers used the study materials for approximately 15 minutes twice each week for 16 weeks. In both conditions, teachers were reminded to use the study materials to supplement, rather than replace, mathematics instruction, and to limit the duration of the activities to 20 minutes per session maximum. Teacher manuals in both conditions provided example scripts teachers could use during the lessons. ICUE lessons included activities, such as using balance scales and recording observations in the workbooks along with solving non-traditional arithmetic problems in the workbook, whereas the active control conditions primarily consisted of students solving non-traditional arithmetic problems in the workbook.

After completing the 32 sessions, teachers administered posttests over five days to minimize student testing fatigue. The first day students completed the same measures as the pretest. The second day, students completed the equation-solving transfer problems. The third day students completed the ITBS items, and the final two days students completed the two MARS tasks, one per day. Teachers were provided with detailed administration instructions and were asked to give all assessments within one week of completing the final session. After the final session, teachers were instructed which eight student workbooks to return.

Analytic model

We used a two-level hierarchical linear model (HLM) to estimate the treatment effect and account for the nested nature of the data (students nested within teachers). The model is shown in equations 1 and 2 below.

$$Posttest_{i_j} = \beta_{0j} + \beta_{1j} Pretest_{ij} + \beta_{2j} Gender_{ij} + \beta_{3j} URG_{ij} + \beta_{4j} EL_{ij} + \epsilon_{ij} \tag{1}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Tx_j + \gamma_{02} MKT_j + \eta_j \tag{2}$$

In this model, the subscripts i and j represent student and teacher, respectively. *Posttest* represents student performance on a posttest measure and *Pretest* represents student performance on a pretest measure. *Gender* (0 = male, 1 = female), *URG* (0 = not a member of an underrepresented racial or ethnic group, 1 = member of an underrepresented racial or ethnic group), and *EL* (0 = English proficient, 1 = English learner) are dichotomous indicator variables for student demographic characteristics as they were recorded in participating districts' student information systems. Consistent with the United States Education Code (2011), we coded students with a reported race or ethnicity of American Indian or Alaskan Native, Black or African-American, Hispanic or Latino, or Pacific Islander as members of underrepresented racial and ethnic groups. *Tx* is a dichotomous variable indicating teachers' condition assignment and *MKT* is teachers' IRT scale score on the MKT assessment. All student pretest and posttest measures were grand-mean-centered, as were teachers' scores on the MKT assessment. The main effect of treatment is captured by γ_{01}. ϵ and η represent random error at the student and teacher levels, respectively. We used the Benjamini-Hochberg method to control for multiple comparisons.

Predictions

We predicted that we would replicate the main findings of the pilot study conducted by McNeil and colleagues (2019a) and that students would make significant gains from pre- to posttest across all measures. As a manipulation check, we investigated the validity of our assumption that the majority of teachers would use study materials as intended with no significant differences between conditions. For the first research question, "Does ICUE lead to improved performance on measures of math equivalence?" we expected that students in the ICUE condition would improve more than students in the active control condition on measures of mathematical equivalence (e.g., encoding, defining the equal sign, equation solving, and MARS tasks). For the second research question, "Do improvements on mathematical equivalence come at the expense of computational fluency?" we did not expect significant differences between conditions on measures of computational fluency (i.e., addition fluency and ITBS).

We included factors for pretest, gender, membership in an underrepresented racial or ethnic group, and English language learner based on their associations with outcomes in prior math education work. Specifically, McNeil et al., (2019a) found that students that use the "add-all" strategy or define the equal sign in incorrect ways are at a disadvantage to learning, and Hornburg, et al., (2017) found that students identifying as female were more likely than students identifying as male to demonstrate these types of errors. More generally, factors of membership in an underrepresented racial or ethnic group or identification as English learner are commonly included to better understand whether whether interventions similarly benefit subpopulations of students (e.g., Bryan et al., 2021; Scammacca, Fall, Capin, Roberts, & Swanson, 2020).

We did not have a priori predictions for our exploratory analysis question, "Do treatment effects interact with participant characteristics?" Our predictions are summarized in Table 2.

Results

Understanding implementation

Overall, both teacher logs and student workbook analyses suggest teachers used the study materials as intended over the course of the study. However, treatment teachers may spent slightly more time preparing for and using the activities.

On the teacher logs, between 75% and 95% of treatment teachers and between 70% and 97% of control teachers reported implementing any given session. On average, treatment teachers reported implementing an average of 27 out of 32 sessions ($SD = 6$) over the course of 29 class periods ($SD = 7$) and 15 weeks ($SD = 4$). Control teachers reported implementing an average of 27 out of 32 sessions ($SD = 5$) over the course of 29 class periods ($SD = 5$) and 14 weeks ($SD = 3$). There were no statistically significant differences between conditions in the number of sessions reported or the total duration of the implementation, whether measured in class periods or weeks.

However, there were differences between conditions in the average amount of time teachers reported preparing for each session and the average duration of each session within a class period. Teachers in both conditions reported a median prep time of 5 minutes per session, but 68% of treatment teachers reported preparing longer than 5 minutes compared to only 53% of control teachers. The difference in prep time between conditions was statistically significant ($W = 1284$, $p < 0.001$). Teachers in the treatment condition reported spending a median of 20 minutes on each session during class, compared to 15 minutes in the control condition. This difference was also statistically significant ($W = 1225$, $p < 0.001$).

As a further implementation check, we requested a stratified random sample of four student workbooks from each teacher to determine whether students participated in each of the 32 supplemental sessions. We collected 222 workbooks from 57 treatment teachers in the analytic sample and 207 workbooks from 57 control teachers in the analytic sample. This represents a teacher response rate of 95% for treatment teachers and 93% for control teachers. For each returned workbook, researchers tallied whether each workbook session had markings indicating that students engaged with the materials.

Eighty-one percent of teachers returned at least one student workbook in which the student had engaged with all 32 sessions. There was relatively low between-teacher variance in workbook completion (intra-class correlation = 0.09) so we used an independent-samples t-test to determine if there were significant differences in workbook completion between the treatment and control groups. The students in the treatment group engaged with an average of 31 sessions ($SD = 2$). The students in the control group engaged with an average of 30 sessions ($SD = 3$). This difference was not statistically significant, $t(315) = -0.428$, $p = 0.669$.

Does the ICUE intervention lead to improved performance on measures of mathematical equivalence, controlling for student demographics and teacher background?

To determine whether the ICUE intervention improved student performance on mathematical equivalence over an active control, we used our model to investigate student performance on the encoding, equal sign naming and defining, equation-solving, and MARS performance task measures. We combined the equation-solving items and equation-solving transfer items in our HLM model.

There were large and statistically significant differences between the conditions on the researcher-developed encoding ($g = 0.19$), equal sign naming and defining ($g = 0.74$), and combined equation-solving measures ($g = 0.78$). The difference in performance between the two conditions on the combined scores for the MARS performance tasks, an independently-developed measure of mathematical equivalence, was also statistically significant ($g = 0.20$).

The unadjusted means for both conditions on each mathematical equivalence measure are shown in Figure 2. The results from the HLM models are shown in Table 3. Performing the Benjamini-Hochberg procedure with the p-values for all treatment effects yields a new alpha threshold of 0.033, meaning that all significant treatment effects on the measures of mathematical equivalence remain statistically significant after correcting for multiple comparisons.

Do improvements in performance on mathematical equivalence come at the expense of computational fluency?

To determine whether improvements in mathematical equivalence came at the expense of computational fluency, we analyzed student performance on measures of arithmetic fluency and the ITBS. No statistically significant differences were found between the treatment and control conditions on either the posttest addition fluency measure or the ITBS. Unadjusted means and standard deviations for both conditions on these measures are shown in Figure 3. See Table 4 for HLM results.

Did student and teacher factors independently predict posttest performance?

The focus of our two research questions was to determine the effects of treatment while controlling for student and teacher factors. Table 3 and Table 4 show the importance of controlling for these factors, as some student and teacher factors predicted posttest performance independent of treatment condition. Specifically, pretest scores were positively correlated with posttest scores on every measure given at both time points. Identifying as female predicted significantly higher posttest performance on the equation encoding and equal sign measures, but not other measures. Membership in an underrepresented group predicted significantly lower posttest performance on all posttest measures except addition fluency. Designation as an English learner predicted significantly lower posttest scores on the equation encoding, equal sign, and ITBS measures. At the teacher level, the teacher Mathematics Knowledge for Teaching score was positively correlated with student posttest performance on MARS tasks.

Do treatment effects interact with participant characteristics?

We explored potential moderators of the treatment effects where confirmatory analyses found a significant treatment effect. We found treatment $\times$ pretest interactions were significant predictors of students' encoding, equation-solving, and MARS task performance (see Table 5). The coefficient of the interaction term is negative, meaning that the treatment effect was strongest for students whose pretest performance was lowest. Performing the Benjamini-Hochberg procedure with the p-values for the treatment effects yields a new alpha threshold of 0.02, meaning that all significant treatment-by-pretest interaction effects on the measures of mathematical equivalence remain statistically significant after correcting for multiple comparisons.

There were no statistically significant interactions between treatment and student demographic characteristics (i.e., gender, membership in an underrepresented racial or ethnic group, or English learner status) or between treatment and teacher MKT performance.

Discussion

The current study found significant positive effects of the ICUE intervention compared with an active control intervention that itself improved learning over business-as-usual practice in a prior study that met What Works Clearinghouse standards without reservations (McNeil, 2014). Children in classrooms that were randomly assigned to the ICUE intervention, which integrates multiple research-based design principles, constructed better understanding of mathematical equivalence than those in classrooms that were randomly assigned to the active control, which included only one research-based design principle, non-traditional arithmetic practice. The findings replicate McNeil et al.'s (2019a) pilot study with a larger and more diverse population of public school teachers and students across the state of California. The prior work was carried out in a small number of classrooms with new first- and second-year teachers enrolled in a Masters program and working in under-resourced parochial schools. The current study found the increased performance on measures of mathematical equivalence were robust, even controlling for multiple factors such as pretest scores, student demographics, and teachers' mathematical knowledge for teaching.

In addition to a larger sample, the current study extends previous work by finding significant differences on an additional transfer measure that was independently created by test developers who are unaffiliated with the developers of the intervention. As described above, the tasks developed by the Mathematics Assessment Resource Service are designed as performance assessments that integrate more complex problem solving and require students to demonstrate conceptual understanding by explaining their processes and procedures. The current study demonstrates that the positive effects of ICUE generalize beyond researcher-created measures.

The study further found no evidence that the benefits of the ICUE intervention came with any detriments to computational fluency. Being able to quickly retrieve basic facts is crucial to freeing up cognitive load, so students can understand explanations and be able to solve problems with multiple steps (e.g., Baroody, Bajwa, & Eiland, 2009). As the ICUE intervention includes activities where students are engaging in balancing a scale or comparing and contrasting different equations and recording their observations in workbooks, students in the active control condition spent proportionally more time solving arithmetic problems, so it was possible that there would be a trade-off between gains in mathematical equivalence versus arithmetic fluency. As children in both conditions improved their performance from pretest to posttest on both the addition fluency measure and the ITBS, and no significant differences were found in fluency growth between the two conditions, we have no reason to think that the ICUE intervention harmed computational fluency while boosting conceptual understanding.

Instructors are often interested in knowing what types of interventions will work for particular subpopulations of students, particularly those that may be struggling (Bryan et al., 2021). Though our sample size design was not intended to be large enough to detect moderation of the treatment effect, exploratory analyses revealed a significant interaction between pretest and treatment, with the ICUE intervention leading to larger gains for students who performed more poorly at pretest. These results are promising as they suggest that the ICUE intervention may be particularly effective for lower performing students. Prior work by McNeil et al., (2019a) found that students that use the "add-all" strategy or define the equal sign in incorrect ways are at a disadvantage to learning. As the current study investigated overall correctness rather than error patterns, future work may be able to discern whether the ICUE intervention helped students successfully overcome operational patterns of thinking more than the active control condition.

Beyond treatment effects, our analyses identified membership in an underrepresented racial or ethnic group, pretest scores, and teacher Mathematics Knowledge for Teaching (MKT) scores as significant predictors of posttest scores. Across all measures pretest scores were positively correlated with posttest scores, identifying as female was positively correlated with performance on equation encoding and equal sign measures, membership in an underrepresented group was negatively correlated with posttest scores on all measures except addition fluency, and designation as an English learner was negatively correlated with performance on measures of encoding, equal sign, and the ITBS. Findings related to membership in an underrepresented racial or ethnic group and pretest scores are generally consistent with student mathematics outcomes on assessments such as the National Assessment of Educational Progress (U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics, 2019).

The emergence of gender (female) as a significant positive predictor of the encoding and equal sign measures (and no relation to the equation solving measures) was unexpected, given prior research that has demonstrated that girls are more likely to solve math equivalence problems incorrectly in the absence of intervention, especially with the traditional “add all” strategy, and no gender differences were observed on encoding and equal sign definition pretest measures in prior work (Hornburg, Rieber, & McNeil, 2017). There is some evidence that girls may be more attuned to their teachers as they show more on-task behavior (Ready, LoGerfo, Burkam, & Lee, 2005), demonstrate greater self discipline (Duckworth & Seligman, 2006), and are more likely to remember mathematical procedures (e.g., Carr & Jessup, 1997; Fennema, Carpenter, Jacobs, Franke, & Levi, 1998; Gallagher et al., 2000). As female students may pay more attention to classroom instruction, in future work we may test a hypothesis that female students may particularly benefit when their teachers explicitly focus on the location and meaning of the equal sign.

The relationship between teachers' MKT scores and posttest scores was consistent with prior findings that connect teacher knowledge of student thinking with achievement outcomes (e.g., Hill, Blunk, et al., 2008; Hill & Chin, 2018). MKT did not predict student performance on either computational fluency measure (e.g. addition fluency or ITBS), and only weakly predicted student performance on the researcher-created equation-solving measure. MKT did predict student achievement on the more conceptual MARS tasks. Though not the primary focus of our study, the finding aligns with the hypothesis that teachers who have a better understanding of what students are thinking are able to tailor instruction in ways that lead to improved outcomes.

Limitations and Future Directions

There were no differences between conditions in the reported number of sessions, weeks of use, or median prep time. However, one limitation of the study was that more treatment teachers than control teachers reported spending over 5 minutes preparing for sessions, and the median reported session time was 20 minutes for treatment teachers compared to 15 minutes for control teachers. Based on the logs, treatment students may have had slightly more time exposed to the intervention. We cannot directly quantify the differences in time as the times were self-reported in five minute increments, and our post-participation survey indicated that some treatment lessons took longer to set up and put away materials (e.g., activities using the balance scales or stickers). Workbook completion data suggest that students in both conditions were exposed to a similar number of lessons.

However, our method of measuring fidelity primarily by examining student work does not consider whether teachers carried out lessons exactly as designed, or whether the quality of classroom discussions differed across conditions. The advantage of an RCT is that it allows us to make broad conclusions about the growth in understanding of mathematical equivalence for students whose teachers administer ICUE versus non-traditional arithmetic practice alone. However, we do not yet fully understand the specific ingredients of ICUE that are driving the observed learning differences. Measuring fidelity differently in future studies of the ICUE materials could help us identify mechanisms contributing to the observed learning differences between conditions.

Another potential limitation could be contamination due to random assignment at the teacher rather than school level. However, we believe the possibility of contamination was minimal for a number of reasons. First, the paper-and-pencil workbooks guided student activities, and teachers were only provided with condition-specific materials. In the treatment condition, each student received an ICUE workbook and teachers received balance scales and stickers for the concreteness fading exercises. In the active control condition, each student received the active control workbook, and teachers were not given any additional manipulatives. Second, participating teachers were provided with detailed teacher guides for each lesson and were asked to avoid co-planning intervention activities with other teachers enrolled in the study. Third, the materials were explicitly designed to be supplementary and not to be used as primary math instruction, so any planning related to the school's math curriculum was not relevant to implementation. Finally, as even minimal contamination of intervention activities would be predicted to *reduce* rather than increase the contrast between treatment and control, the substantial differences reported in this paper would underestimate the true impact of the treatment intervention.

The current study demonstrates the effectiveness of the ICUE intervention with a large-scale cluster-randomized trial, but future work is needed to understand the nuances and long-term consequences of these effects. For example, findings reveal that the comprehensive ICUE intervention, which includes three components beyond non-traditional arithmetic practice, leads children to construct better understanding of mathematical equivalence when compared to spending the full intervention time on non-traditional arithmetic practice alone. This is important because the non-traditional arithmetic practice alone had already been shown to improve children's understanding in a study that met WWC guidelines without reservations (McNeil et al., 2011). Before this study, some may have argued that intervention time would be better spent focusing exclusively on the activity that had already been shown to work. Although we can use our results to conclude that ICUE is better than non-traditional arithmetic alone, our study does not allow us to tease out the independent effects of each of the additional three components, so that remains a question for future work.

As the posttest in the current study was administered within a week of completing the intervention, we can only conclude that students demonstrate better immediate learning outcomes after engaging in ICUE versus nontraditional arithmetic practice alone. A previous study showed that the benefits of nontraditional arithmetic practice over traditional practice last for five to six months (McNeil et al., 2015), so it will be important to follow students to determine if the ICUE intervention keeps its advantage over nontraditional arithmetic across time. Additionally, the sample of students in the present study was representative of the population in California, but additional studies may seek to replicate these findings with broader populations of elementary students across the United States.

Finally, the ICUE intervention included four research-based design principles. Since the initial development of the intervention, other promising factors have been identified, and incorporating these components may improve student learning even further. For example, recent work shows benefits of students' metacognitive reflection during math equivalence lessons (Fyfe, Byers, & Nelson, 2021) and other work highlights the importance of student prior knowledge as a moderator in the benefits of feedback (Fyfe & Brown, 2018; Fyfe & Rittle-Johnson, 2016). Future work may explore whether and how integrating additional components may future improve student learning. In particular, ongoing work is exploring whether and how tailored feedback in an interactive version of the workbook may help students more readily notice errors and learn correct conceptions.

Conclusion

Students' understanding of mathematical equivalence is included in state standards such as the Common Core (National Governors Association et al., 2010), but most elementary instruction lacks research-based practices to develop this knowledge (Silla, Hornburg, Kloser, & McNeil, 2020). The current study provides teachers and policymakers with clear evidence that appropriate modifications to arithmetic practice in early elementary school can help children construct a formal understanding of mathematical equivalence while also building computational fluency (Charles, 2005). The current study demonstrates that the ICUE intervention improves children's understanding of mathematical equivalence early in elementary school, which evidence suggests will provide a critical foundation for later success in algebra (Hornburg et al., 2022).

References

Adelman, C. (2006). The toolbox revisited: Paths to degree completion from high school through college. *US Department of Education.*

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology*, *35*(1), 127.

Alibali, M. W., Knuth, E. J., Hattikudur, S., McNeil, N. M., & Stephens, A. C. (2007). A longitudinal examination of middle school students' understanding of the equal sign and equivalent equations. *Mathematical Thinking and Learning*, *9*(3), 221–247. doi: 10.1080/10986060701360902

Alibali, M. W., Phillips, K. M., & Fischer, A. D. (2009). Learning new problem-solving strategies leads to changes in problem representation. *Cognitive Development*, *24*(2), 89–101. doi: 10.1016/j.cogdev.2008.12.005

Baroody, A. J., Bajwa, N. P., & Eiland, M. (2009). Why can't johnny remember the basic facts? *Developmental Disabilities Research Reviews*, *15*(1), 69–79.

Baroody, A. J., & Ginsburg, H. P. (1983). The effects of instruction on children's understanding of the "equals" sign. *The Elementary School Journal*, *84*(2), 199–212. doi: 10.1086/461356

Blanton, M., & Kaput, J. (2005). Characterizing a classroom practice that promotes algebraic reasoning. *Journal for Research in Mathematics Education*, *36*(5), 412–446. doi: 10.2307/30034944

Blanton, M., Stephens, A., Knuth, E., Gardiner, A. M., Isler, I., & Kim, J.-S. (2015). The development of children's algebraic thinking: The impact of a comprehensive early algebra intervention in third grade. *Journal for Research in Mathematics Education*, *46*(1), 39–87. doi: 10.5951/jresematheduc.46.1.0039

Blanton, M., Stroud, R., Stephens, A., Gardiner, A. M., Stylianou, D. A., Knuth, E., ... Strachota, S. (2019). Does early algebra matter? the effectiveness of an early algebra intervention in grades 3 to 5. *American Educational Research Journal*, *56*(5), 1930–1972. doi: 10.3102/0002831219832301

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, *64*(2), 123. doi: https://doi.org/10.1037/h0043805

Bruner, J. S. (1966). *Toward a theory of instruction.* Cambridge, MA: Belknap Press.

Bryan, C. J., Tipton, E., & Yeager, D. S. (2021). Behavioural science is unlikely to change the world without a heterogeneity revolution. *Nature Human Behaviour*, *5*(8), 980–989.

Byrd, C. E., McNeil, N. M., Chesney, D. L., & Matthews, P. G. (2015). A specific misconception of the equal sign acts as a barrier to children's learning of early algebra. *Learning and Individual Differences*, *38*, 61–67. doi: 10.1016/j.lindif.2015.01.001

California Department of Education. (n.d.). *State performance overview: State of California.* Retrieved 2021-11-16, from `https://www.caschooldashboard.org/reports/ca/2020`

California Department of Education. (2020a). *Facts about English learners in California - CalEdFacts.* Retrieved 2020-11-06, from `https://www.cde.ca.gov/ds/sd/cb/cefelfacts.asp`

California Department of Education. (2020b). *Fingertip facts on education in California - CalEdFacts.* Retrieved 2020-11-06, from `https://www.cde.ca.gov/ds/sd/cb/ceffingertipfacts.asp`

Carpenter, T. P., Franke, M. L., & Levi, L. (2003). *Thinking mathematically: Integrating arithmetic and algebra in elementary school.* Portsmouth, NH: Heinemann.

Carpenter, T. P., & Levi, L. (2000). *Developing conceptions of algebraic reasoning in the primary grades.* Madison, WI: National Center for Improving Student Learning and Achievement in Mathematics and Science.

Carr, M., & Jessup, D. L. (1997). Gender differences in first-grade mathematics strategy use: Social and metacognitive influences. *Journal of Educational Psychology*, *89*(2), 318.

Charles, R. I. (2005). Big ideas and understandings as the foundation for elementary and middle school mathematics. *National Council for Supervisors of Mathematics: Journal of Mathematics Education Leadership*, *8*(2), 9–24.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55–81. doi: 10.1016/0010-0285(73)90004-2

Chesney, D. L., McNeil, N. M., Petersen, L. A., & Dunwiddie, A. E. (2018). Arithmetic practice that includes relational words promotes understanding of symbolic equations. *Learning and Individual Differences*, *64*, 104–112.

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*(2), 1047–1058. doi: 10.1016/j.cognition.2007.04.010

Darling-Hammond, L., & Falk, B. (2013). *Teacher learning through assessment: How student performance assessments can support teacher learning* (Tech. Rep.). Center for American Progress. Retrieved from https://www.americanprogress.org/wp-content/uploads/2013/09/TeacherLearning.pdf

Denmark, T., Barco, E., & Voran, J. (1976). Final report: A teaching experiment on equality (pmdc technical report no. 6). *Tallahassee: Florida State University.*

Duckworth, A. L., & Seligman, M. E. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology*, *98*(1), 198.

Falkner, K. P., Levi, L., & Carpenter, T. P. (1999). Children's understanding of equality: A foundation for algebra. *Teaching Children Mathematics*, *6*(4), 232–236.

Fennema, E., Carpenter, T. P., Jacobs, V. R., Franke, M. L., & Levi, L. W. (1998). A longitudinal study of gender differences in young children's mathematical thinking. *Educational Researcher*, *27*(5), 6–11.

Foster, D., & Noyce, P. (2004). The mathematics assessment collaborative: Performance testing to improve instruction. *Phi Delta Kappan*, *85*(5), 367-374. doi:

10.1177/003172170408500507

Fyfe, E. R., & Brown, S. A. (2018). Feedback influences children's reasoning about math equivalence: A meta-analytic review. *Thinking & Reasoning*, *24*(2), 157–178.

Fyfe, E. R., Byers, C., & Nelson, L. J. (2021). The benefits of a metacognitive lesson on children's understanding of mathematical equivalence, arithmetic, and place value. *Journal of Educational Psychology*.

Fyfe, E. R., Matthews, P. G., & Amsel, E. (2020). College developmental math students' knowledge of the equal sign. *Educational Studies in Mathematics*, *104*(1), 65–85.

Fyfe, E. R., Matthews, P. G., Amsel, E., McEldoon, K. L., & McNeil, N. M. (2018). Assessing formal knowledge of math equivalence among algebra and pre-algebra students. *Journal of Educational Psychology*, *110*(1), 87–101. Retrieved from `https://doi.org/10.1037/edu0000208` doi: 10.1037/edu0000208

Fyfe, E. R., McNeil, N. M., & Borjas, S. (2015). Benefits of "concreteness fading" for children's mathematics understanding. *Learning and Instruction*, *35*, 104–120. Retrieved from `http://dx.doi.org/10.1016/j.learninstruc.2014.10.004` doi: 10.1016/j.learninstruc.2014.10.004

Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review*, *26*(1), 9–25.

Fyfe, E. R., & Rittle-Johnson, B. (2016). Feedback both helps and hinders learning: The causal role of prior knowledge. *Journal of Educational Psychology*, *108*(1), 82.

Gallagher, A. M., De Lisi, R., Holst, P. C., McGillicuddy-De Lisi, A. V., Morely, M., & Cahalan, C. (2000). Gender differences in advanced mathematical problem solving. *Journal of Experimental Child Psychology*, *75*(3), 165–190.

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *The Journal of the Learning Sciences*, *14*(1), 69–110. Retrieved from `https://doi.org/10.1207/s15327809jls1401_4` doi:

10.1207/s15327809jls1401\{_}4

Gravemeijer, K. (2002). Preamble: From models to modeling. In K. Gravemeijer, R. Lehrer, B. Oers, & L. Verschaffel (Eds.), *Symbolizing, modeling and tool use in mathematics education* (pp. 7–22). Dordrecht, The Netherlands: Springer.

Hattikudur, S., & Alibali, M. W. (2010). Learning about the equal sign: Does comparing with inequality symbols help? *Journal of Experimental Child Psychology*, *107*(1), 15–30. Retrieved from https://doi.org/10.1016/j.jecp.2010.03.004 doi: 10.1016/j.jecp.2010.03.004

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking "pedagogical content knowledge": Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, *39*(4), 372–400. Retrieved from https://www.jstor.org/stable/40539304

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, *26*(4), 430–511.

Hill, H. C., & Chin, M. (2018). Connections between teachers' knowledge of students, instruction, and achievement outcomes. *American Educational Research Journal*, *55*(5), 1076–1112.

Hornburg, C. B. (2017). *Optimizing problem format to facilitate children's understanding of math equivalence* (Unpublished doctoral dissertation). University of Notre Dame.

Hornburg, C. B., Brletic-Shipley, H., Matthews, J. M., & McNeil, N. M. (2021). Improving understanding of mathematical equivalence. *Mathematics Teacher: Learning and Teaching PK-12*, *114*(1), 16–26. Retrieved from https://doi.org/10.5951/MTLT.2020.0109

Hornburg, C. B., Devlin, B. L., & McNeil, N. M. (2022). Earlier understanding of mathematical equivalence in elementary school predicts greater algebra readiness in middle school. *Journal of Educational Psychology*, *114* (3), 540–559.

Hornburg, C. B., Rieber, M. L., & McNeil, N. M. (2017). An integrative data analysis of gender differences in children's understanding of mathematical equivalence. *Journal of Experimental Child Psychology*, *163*, 140–150.

Hussar, B., Zhang, J., Hein, S., Wang, K., Roberts, A., Cui, J., … Dilig, R. (2020). *The condition of education 2020. nces 2020-144.* US Department of Education, National Center for Education Statistics. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2020144.rev

Jacobs, V. R., Franke, M. L., Carpenter, T. P., Levi, L., & Battey, D. (2007). Professional development focused on children's algebraic reasoning in elementary school. *Journal for Research in Mathematics Education*, *38* (3), 258–288.

Jones, I., Inglis, M., Gilmore, C., & Dowens, M. (2012). Substitution and sameness: Two components of a relational conception of the equals sign. *Journal of Experimental Child Psychology*, *113* (1), 166–176. Retrieved from http://dx.doi.org/10.1016/j.jecp.2012.05.003 doi: 10.1016/j.jecp.2012.05.003

Knuth, E. J., Stephens, A. C., McNeil, N. M., & Alibali, M. W. (2006). Does understanding the equal sign matter? evidence from solving equations. *Journal for Research in Mathematics Education*, *37* (4), 297–312.

Li, X., Ding, M., Capraro, M. M., & Capraro, R. M. (2008). Sources of differences in children's understandings of mathematical equality: Comparative analysis of teacher guides and student texts in china and the united states. *Cognition and Instruction*, *26* (2), 195–217. doi: 10.1080/07370000801980845

Loveless, T. (2008). The misplaced math student: Lost in eighth-grade algebra. the 2008 brown center report on american education. special release. *Brookings Institution*.

Ma, X. (2001). A longitudinal assessment of antecedent course work in mathematics and subsequent mathematical attainment. *The Journal of Educational Research*, *94* (1), 16–28. doi: 10.1080/00220670009598739

Mathematics Assessment Resource Service. (2004). *Tools for teachers: Grade 2.* Retrieved from http://www.svmimac.org

Mathematics Assessment Resource Service. (2007). *Tools for teachers: Grade 2, part 3.* Retrieved from http://www.svmimac.org

Matthews, P. G., & Fuchs, L. S. (2020). Keys to the gate? equal sign knowledge at second grade predicts fourth-grade algebra competence. *Child Development*, *91*(1), 14–28.

McNeil, N. M. (2007). U-shaped development in math: 7-year-olds outperform 9-year-olds on equivalence problems. *Developmental Psychology*, *43*(3), 687–695. doi: 10.1037/0012-1649.43.3.687

McNeil, N. M. (2008). Limitations to teaching children 2+ 2= 4: Typical arithmetic problems can hinder learning of mathematical equivalence. *Child Development*, *79*(5), 1524–1537. doi: 10.1111/j.1467-8624.2008.01203.x

McNeil, N. M. (2014). A "change–resistance" account of children's difficulties understanding mathematical equivalence. *Child Development Perspectives*, *8*(1), 42–47. doi: 10.1111/cdep.12062

McNeil, N. M., & Alibali, M. W. (2000). Learning mathematics from procedural instruction: Externally imposed goals influence what is learned. *Journal of Educational Psychology*, *92*(4), 734–744. doi: 10.1037/0022-0663.92.4.734

McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science*, *28*(3), 451–466.

McNeil, N. M., & Alibali, M. W. (2005a). Knowledge change as a function of mathematics experience: All contexts are not created equal. *Journal of Cognition and Development*, *6*(2), 285–306. doi: 10.1207/s15327647jcd0602\{_}âĂİ6

McNeil, N. M., & Alibali, M. W. (2005b). Why won't you change your mind? knowledge of operational patterns hinders learning and performance on equations. *Child Development*, *76*(4), 883–899. doi: 10.1111/j.1467-8624.2005.00884.x

McNeil, N. M., Chesney, D. L., Matthews, P. G., Fyfe, E. R., Petersen, L. A., Dunwiddie, A. E., & Wheeler, M. C. (2012). It pays to be organized: Organizing arithmetic practice around equivalent values facilitates understanding of math equivalence. *Journal of Educational Psychology*, *104*(4), 1109–1121. doi: 10.1037/a0028997

McNeil, N. M., Fyfe, E. R., & Dunwiddie, A. E. (2015). Arithmetic practice can be modified to promote understanding of mathematical equivalence. *Journal of Educational Psychology*, *107*(2), 423–436. doi: 10.1037/a0037687

McNeil, N. M., Fyfe, E. R., Petersen, L. A., Dunwiddie, A. E., & Brletic-Shipley, H. (2011). Benefits of practicing 4= 2+ 2: Nontraditional problem formats facilitate children's understanding of mathematical equivalence. *Child Development*, *82*(5), 1620–1633. doi: 10.1111/j.1467-8624.2011.01622.x

McNeil, N. M., Grandau, L., Knuth, E. J., Alibali, M. W., Stephens, A. C., Hattikudur, S., & Krill, D. E. (2006). Middle-school students' understanding of the equal sign: The books they read can't help. *Cognition and Instruction*, *24*(3), 367–385. doi: 10.1207/s1532690xci2403_3

McNeil, N. M., Hornburg, C. B., Brletic-Shipley, H., & Matthews, J. M. (2019a). Improving children's understanding of mathematical equivalence via an intervention that goes beyond nontraditional arithmetic practice. *Journal of Educational Psychology*, *111*(6), 1023–1044. doi: 10.1037/edu0000337

McNeil, N. M., Hornburg, C. B., Devlin, B. L., Carrazza, C., & McKeever, M. O. (2019b). Consequences of individual differences in children's formal understanding of mathematical equivalence. *Child Development*, *90*(3), 940–956. doi: 10.1111/cdev.12948

McNeil, N. M., Hornburg, C. B., Fuhs, M. W., & Connor, D. (2017). Understanding children's difficulties with mathematical equivalence. In *Acquisition of complex arithmetic skills and higher-order mathematics concepts* (pp. 167–195). Elsevier. doi: 10.1016/B978-0-12-805086-6.00008-4

Mickey, K. W., & McClelland, J. L. (2014). A neural network model of learning mathematical equivalence. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 36). Retrieved from `https://escholarship.org/content/qt3xq6j3cs/qt3xq6j3cs.pdf`

National Governors Association et al. (2010). Common core state standards. *Washington, DC*. `http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf`.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the national mathematics advisory panel*. Washington, DC: US Department of Education.

Perry, M. (1991). Learning and transfer: Instructional conditions and conceptual change. *Cognitive Development*, *6*(4), 449–468. doi: 10.1016/0885-2014(91)90049-J

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development*, *3*(4), 359–400. doi: 10.1016/0885-2014(88)90021-4

Powell, S. R. (2012). Equations and the equal sign in elementary mathematics textbooks. *The Elementary School Journal*, *112*(4), 627–648.

Powell, S. R., & Fuchs, L. S. (2010). Contribution of equal-sign instruction beyond word-problem tutoring for third-grade students with mathematics difficulty. *Journal of Educational Psychology*, *102*(2), 381–394. doi: 10.1037/a0018447

Ready, D. D., LoGerfo, L. F., Burkam, D. T., & Lee, V. E. (2005). Explaining girls' advantage in kindergarten literacy learning: Do classroom behaviors make a difference? *The Elementary School Journal*, *106*(1), 21–38.

Renwick, E. (1932). Children's misconceptions concerning the symbols for mathematical equality. *British Journal of Educational Psychology*, *2*(2), 173–183. doi: 10.1111/j.2044-8279.1932.tb02743.x

Rittle-Johnson, B. (2006). Promoting transfer: The effects of direct instruction and self-explanation. *Child Development*, *77*(1), 1–15. doi:

10.1111/j.1467-8624.2006.00852.x

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology*, *91*(1), 175–189. doi: 10.1037/0022-0663.91.1.175

Scammacca, N., Fall, A.-M., Capin, P., Roberts, G., & Swanson, E. (2020). Examining factors affecting reading and math growth and achievement gaps in grades 1–5: A cohort-sequential longitudinal approach. *Journal of Educational Psychology*, *112*(4), 718.

Schliemann, A. D. (2007). *Bringing out the algebraic character of arithmetic: From children's ideas to classroom practice*. Lawrence Erlbaum Associates.

Seidenberg, M. S., & Zevin, J. D. (2006). Connectionist models in developmental cognitive neuroscience: Critical periods and the paradox of success. *Processes of Change in Brain and Cognitive development: Attention and Performance XXI*, 315–347.

Sherman, J., & Bisanz, J. (2009). Equivalence in symbolic and nonsymbolic contexts: Benefits of solving problems with manipulatives. *Journal of Educational Psychology*, *101*(1), 88-100. doi: 10.1037/a0013156

Siegler, R. S. (2002). Microgenetic studies of self-explanation. In *Microdevelopment: Transition processes in development and learning* (pp. 31–58). New York: Cambridge University Press.

Silla, E. M., Hornburg, C. B., Kloser, M., & McNeil, N. M. (2020). Research-based teaching practices for improving students' understanding of mathematical equivalence have not made it into elementary classrooms. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd annual conference of the cognitive science society* (pp. 2937–2943). Cognitive Science Society.

United States Code. (2011). *20 U.S.C. §1067k. Definitions.* Retrieved from `https://www.govinfo.gov/content/pkg/USCODE-2011-title20/html/USCODE-2011-title20-chap28-subchapIII-partE-subpart3-sec1067k.htm`

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2019). *National assessment of educational progress, 2019 mathematics assessment.* Retrieved from https://www.nationsreportcard.gov/mathematics

What Works Clearinghouse, Institute of Education Sciences. (2012). *Elementary school mathematics evidence review protocol v2.0.* Retrieved from https://ies.ed.gov/ncee/wwc/Document/21

Table 1

Demographics of students in the analytic sample

Demographic	% in Control	% in Treatment	% in Study	% in CA
African American, not Hispanic*	4.9%	5.7%	5.3%	5.3%
American Indian or Alaska Native*	0.6%	0.6%	0.6%	0.5%
Asian	8.2%	8.8%	8.5%	9.3%
Filipino	2.1%	2.7%	2.4%	2.4%
Hispanic or Latino*	43.3%	41.8%	42.5%	54.9%
Pacific Islander*	0.9%	1.5%	1.2%	0.4%
White, not Hispanic	32.4%	30.9%	31.6%	22.4%
Two or more races, not Hispanic	7.6%	8.1%	7.9%	3.9%
Ethnicity not reported*	NA	NA	0.0%	0.9%
English learner*	31.8%	30.2%	31.0%	18.6%

Note:

* These racial/ethnic groups are coded as underrepresented groups.

Table 2

Predicted pretest to posttest gains in student performance by measure and experimental condition

Measure	Construct	Treatment	Active Control
	Researcher-developed measures		
Equation-solving	Mathematical equivalence	Large gain	Gain
Addition fluency	Computational fluency	Gain	Gain
	Independently-developed measures		
MARS tasks	Mathematical equivalence	Large gain	Gain
ITBS	Mathematics achievement	Gain	Gain

Table 3

HLM to estimate effects of covariates on mathematical equivalence measures

	B	SE	df	t	p	Sig.
Equation Encoding						
(Intercept)	-0.06	0.05	207.76	-1.03	0.304	
Treatment	0.19	0.06	113.46	3.02	0.003	**
Teacher MKT score	0.06	0.03	117.37	1.85	0.067	+
Pre-test	0.31	0.02	2251.58	15.57	< 0.001	***
Female	0.18	0.04	2172.15	4.89	< 0.001	***
Underrepresented group	-0.20	0.05	1821.68	-4.29	< 0.001	***
English learner	-0.12	0.05	2209.03	-2.46	0.014	*
Equal Sign Naming + Defining						
(Intercept)	-0.10	0.06	189.29	-1.61	0.109	
Treatment	0.28	0.08	117.96	3.63	< 0.001	***
Teacher MKT score	0.00	0.04	120.67	0.04	0.971	
Pre-test	0.24	0.02	2234.17	12.25	< 0.001	***
Female	0.12	0.04	2162.00	3.36	0.001	***
Underrepresented group	-0.15	0.05	2079.31	-3.07	0.002	**
English learner	-0.10	0.05	2251.47	-2.11	0.035	*
Equation Solving: Combined						
(Intercept)	-0.19	0.05	197.39	-3.61	< 0.001	***
Treatment	0.73	0.06	113.14	11.65	< 0.001	***
Teacher MKT score	0.06	0.03	116.41	1.81	0.073	+
Pre-test	0.33	0.02	2250.22	18.21	< 0.001	***
Female	-0.02	0.03	2163.53	-0.54	0.59	
Underrepresented group	-0.30	0.04	1948.36	-6.88	< 0.001	***
English learner	-0.07	0.04	2236.25	-1.64	0.101	
MARS Tasks						
(Intercept)	0.06	0.06	185.38	1.04	0.298	
Treatment	0.20	0.07	111.36	2.93	0.004	**
Teacher MKT score	0.09	0.03	114.11	2.63	0.01	**
Addition fluency pre-test	0.32	0.02	2250.47	16.71	< 0.001	***
Equation-solving pre-test	0.24	0.02	2245.62	13.19	< 0.001	***
Female	-0.01	0.03	2155.36	-0.20	0.845	
Underrepresented group	-0.27	0.04	2077.88	-6.24	< 0.001	***
English learner	-0.08	0.04	2248.83	-1.94	0.052	+

Note:

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4

HLM to estimate effects of predictors on measures of computational fluency

	B	SE	df	t	p	Sig.
			Addition Fluency			
(Intercept)	-0.05	0.10	126.34	-0.46	0.644	
Treatment	0.09	0.14	117.60	0.63	0.528	
Teacher MKT score	-0.01	0.07	117.90	-0.12	0.903	
Pre-test	0.38	0.01	2162.66	31.11	< 0.001	***
Female	-0.02	0.02	2136.59	-0.99	0.32	
Underrepresented group	0.01	0.03	2186.80	0.39	0.7	
English learner	0.02	0.03	2159.78	0.63	0.528	
(Intercept)	0.15	0.06	171.85	2.61	0.01	**
		ITBS				
Treatment	0.05	0.07	110.05	0.75	0.453	
Teacher MKT score	0.03	0.04	112.30	0.98	0.327	
Addition fluency pre-test	0.31	0.02	2249.41	16.97	< 0.001	***
Equation-solving pre-test	0.25	0.02	2235.45	14.62	< 0.001	***
Female	0.01	0.03	2149.20	0.34	0.735	
Underrepresented group	-0.27	0.04	2156.31	-6.58	< 0.001	***
English learner	-0.16	0.04	2250.51	-4.04	< 0.001	***

Note:

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5

HLM to estimate treatment x pre-test interaction effects

	B	SE	df	t	p	Sig.
Equation Encoding						
(Intercept)	-0.06	0.05	206.86	-1.03	0.304	
Treatment	0.19	0.06	113.35	3.00	0.003	**
Teacher MKT score	0.06	0.03	117.21	1.83	0.071	+
Pre-test	0.35	0.03	2250.54	12.37	< 0.001	***
Female	0.18	0.04	2170.73	4.92	< 0.001	***
Underrepresented group	-0.20	0.05	1828.88	-4.26	< 0.001	***
English learner	-0.12	0.05	2209.74	-2.43	0.015	*
Treatment x pretest	-0.09	0.04	2245.33	-2.33	0.02	*
Equation-Solving: Combined						
(Intercept)	-0.19	0.05	197.01	-3.78	< 0.001	***
Treatment	0.73	0.06	113.25	11.77	< 0.001	***
Teacher MKT score	0.05	0.03	116.61	1.64	0.103	
Pre-test	0.48	0.02	2250.75	19.46	< 0.001	***
Female	-0.01	0.03	2162.80	-0.40	0.686	
Underrepresented group	-0.29	0.04	1953.57	-6.88	< 0.001	***
English learner	-0.07	0.04	2236.25	-1.68	0.094	+
Treatment x pretest	-0.31	0.03	2241.54	-8.95	< 0.001	***
MARS Tasks						
(Intercept)	0.06	0.06	184.06	1.01	0.312	
Treatment	0.20	0.07	111.17	2.89	0.005	**
Teacher MKT score	0.09	0.03	113.96	2.55	0.012	*
Addition fluency pre-test	0.35	0.03	2248.67	12.39	< 0.001	***
Equation-solving pre-test	0.29	0.03	2238.67	10.88	< 0.001	***
Female	0.00	0.03	2152.99	-0.13	0.894	
Underrepresented group	-0.27	0.04	2084.95	-6.23	< 0.001	***
English learner	-0.08	0.04	2247.46	-1.95	0.051	+
Treatment x addition fluency	-0.06	0.04	2245.69	-1.50	0.135	
Treatment x equation solving	-0.09	0.04	2245.35	-2.38	0.017	*

Note:

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(a) *ICUE intervention materials* (b) *Active control materials*

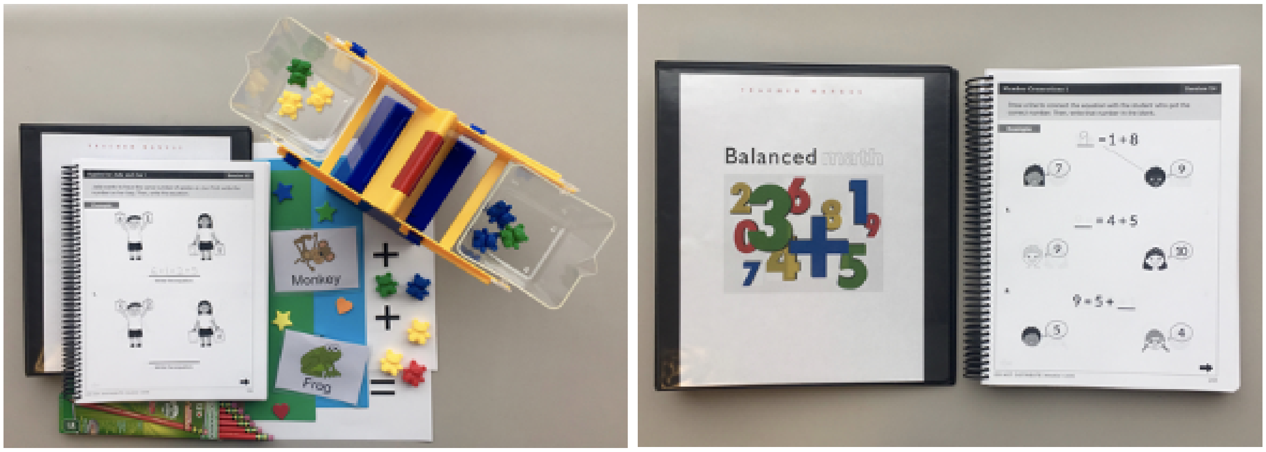

Figure 1

Curriculum materials for the two study conditions.

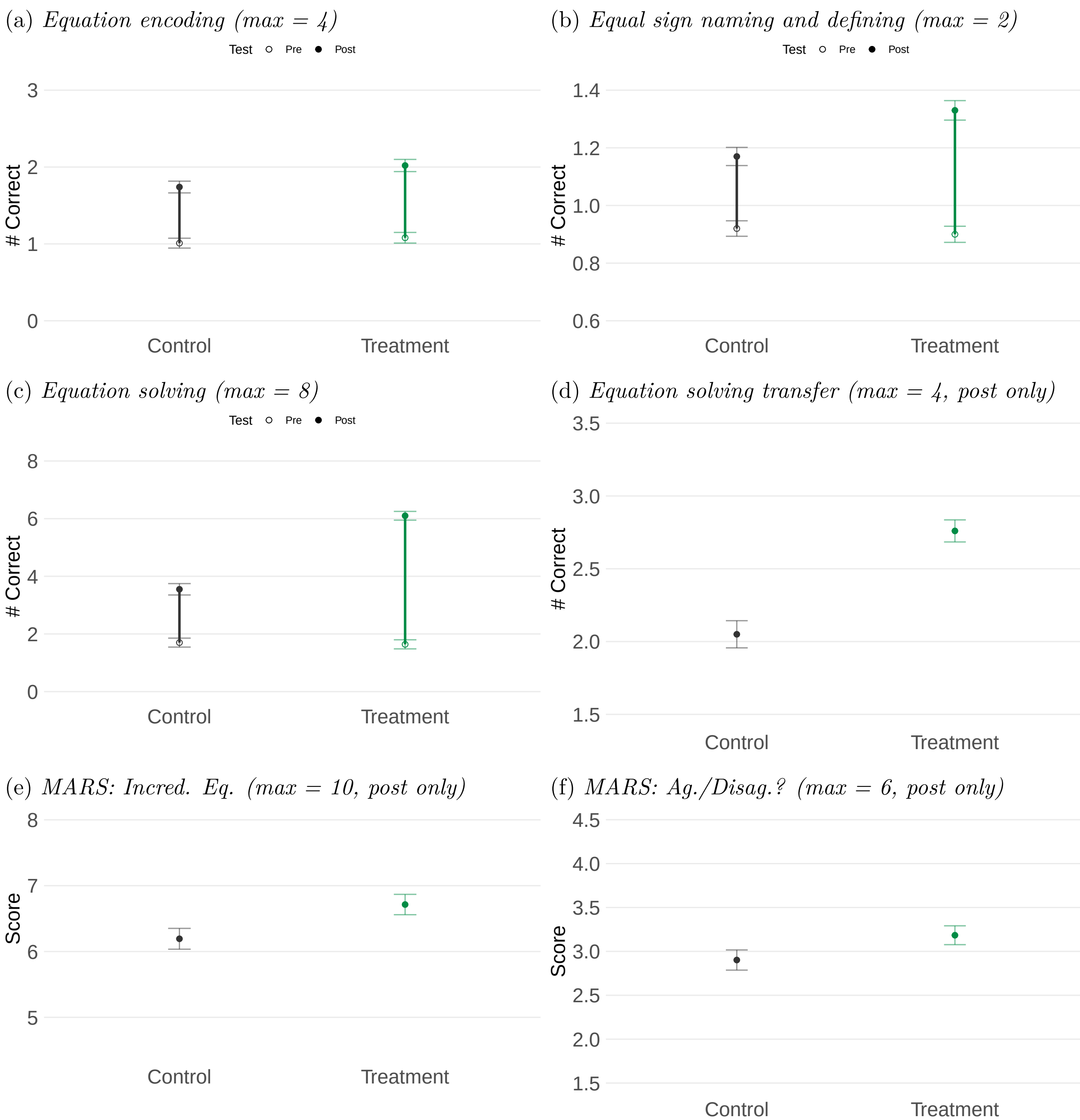

Figure 2

Unadjusted mean student performance on mathematical equivalency measures. Error bars represent 95% confidence intervals.

(a) *Addition fluency (max = 81)* (b) *ITBS (max = 31)*

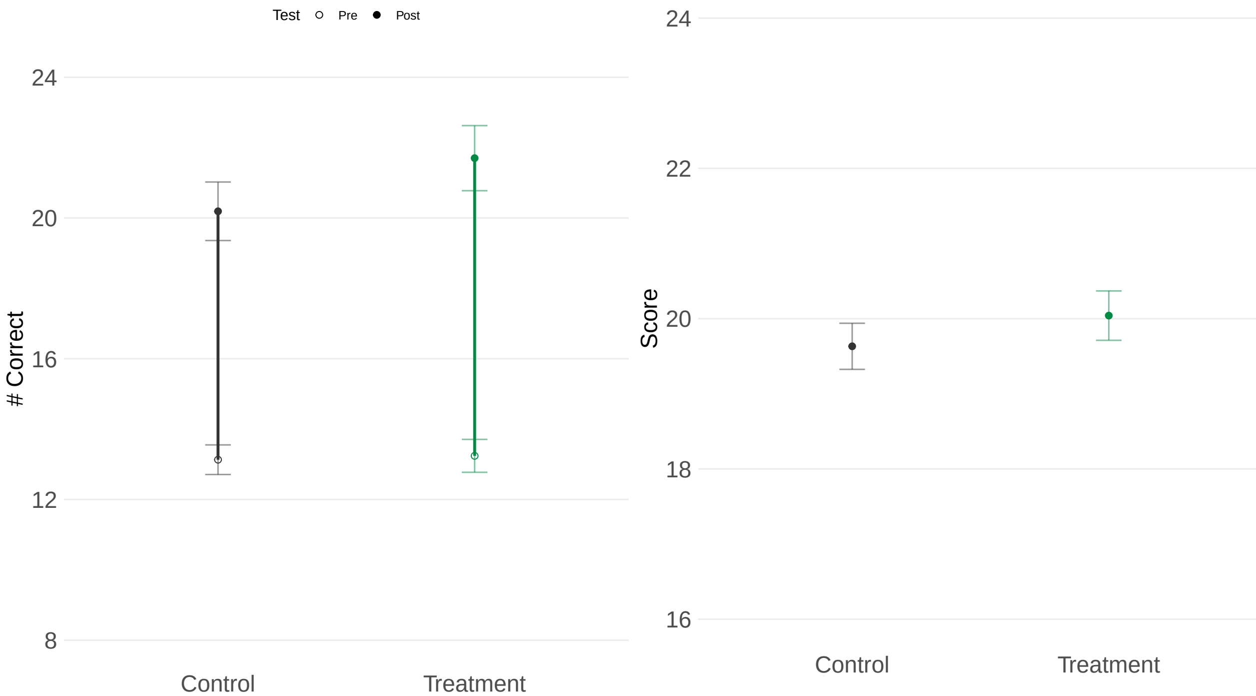

Figure 3

Unadjusted mean student performance on computational fluency measures. Error bars represent 95% confidence intervals.