Predicting Learning and Knowledge Transfer in Two Early Mathematical Equivalence Interventions

Kristen Johannes (kjohann@WestEd.org),
WestEd, 2470 Mariner Square Loop, Alameda, CA 94501 USA

Nicole McNeil (nmcneil@nd.edu)
University of Notre Dame, 102 Haggar Hall, Notre Dame, IN 46556 USA

Yvonne Kao (ykao@WestEd.org)
WestEd, 730 Harrison Street, San Francisco, CA 94107 USA

Jodi Davenport (jdavenp@WestEd.org)
WestEd, 2470 Mariner Square Loop, Alameda, CA 94501 USA

Abstract

Many students fail to develop adequate understanding of mathematical equivalence in early grades, which impacts later algebra learning. Work from McNeil and colleagues proposes that this failure is partly due to the format of traditional instruction and practice with highly similar problems, which encourages students to develop ineffective representations of problem types (McNeil, 2014, McNeil & Alibali, 2005). In the current study, we explore students' learning trajectories in two matched equivalence interventions. We show that, relative to an active control, the principle-based treatment intervention gives rise to a greater number of successful learners, a designation that, in turn, leads to improved performance on distal transfer assessments. We further demonstrate a predictive relationship between students' engagement with the intervention, via workbook completion, and likelihood of becoming a successful learner. Our findings have implications for early detection of learning and subsequent scaffolding for low-performing students.

Keywords: Mathematical representations; learning; mathematics education; randomized control trial

Introduction

What kind of early intervention can best help students learn key concepts and prevent later struggles in algebra? Research suggests that understanding mathematical equivalence is a critical component of algebraic reasoning (Carpenter, Franke, & Levi, 2003; Charles, 2005; Knuth, Stephens, McNeil, & Alibali, 2006). However, the majority of US students fail to reason with and apply concepts of equivalence (McNeil & Alibali, 2005), making encoding errors when remembering mathematical equations (e.g., McNeil & Alibali, 2004), and interpreting the equal sign to mean "calculate the total" rather than "two amounts are the same" (e.g., Behr, Erlwanger, & Nichols, 1980).

McNeil and Alibali (2005) proposed a change-resistance account of children's difficulty with equivalence: traditional arithmetic instruction that focuses on procedures (i.e., solving problems such as 3 + 4 = _) promotes a misconception of the equal sign as a request for an answer and interferes with the development of relational understanding. The majority of examples of arithmetic problems in early elementary math curricula show operations (e.g., addition and subtraction) on the left of the equal sign and the "answer" on the right (Seo & Ginsburg, 2003). Based on this observation, McNeil and Alibali characterize the representations that develop in early mathematics as "operational patterns" as they reflect an understanding of arithmetic that focuses on the operators (e.g., +, –, ×, ÷) rather than the relational nature of mathematical equations. Once entrenched, children rely on these potentially misleading patterns when encoding, interpreting, and solving novel mathematics problems. Students who expect all problems to have operations on the left fail to correctly encode the problem being asked. For instance, after briefly viewing the problem "7 + 4 + 5 = 7 + __" many children rely on their knowledge of an "operations = answer" problem format and erroneously remember the problem as "7 + 4 + 5 + 7 = __" (McNeil & Alibali, 2004). Students also struggle to interpret what a mathematical problem is asking. When asked to define the equal sign—even in the context of a mathematical equivalence problem—many children treat it like an arithmetic operator (+ or -) that means they should calculate the total (McNeil & Alibali, 2005). Finally, entrenched patterns mislead students to solve the problem "7 + 4 + 5 = 7 + __" by performing all given operations on all given numbers and put 23 (instead of 9) in the blank (McNeil, 2007; Rittle-Johnson, 2006). These findings support the idea that children's difficulties with mathematical equivalence are partially due to inappropriate knowledge derived from overly narrow experience with traditional arithmetic.

ICUE: Improving Children's Understanding of Equivalence Intervention

As current math practice seems to promote the development of faulty representations, the change resistance account of "operational patterns" offers design principles for instruction to improve students' understanding of equivalence. Initially, researchers hypothesized that greater exposure to "non-traditional arithmetic" problems (e.g., presenting operations on the right side of the equation, "_ = 2 + 4" and using relational phrases such as "is equal to" instead of the equal sign in practice problems) may prevent students from developing operational patterns (McNeil et al., 2011). Though practice with non-traditional arithmetic led to improved outcomes over traditional instruction, a number of students failed to reach proficiency (McNeil, Fyfe, & Dunwiddle, 2015).

To further promote mastery of equivalence, McNeil and colleagues added additional design features beyond non-traditional arithmetic practice. The current version of the materials, dubbed Improving Children's Understanding of Equivalence (ICUE), consists of second grade student activities that reduce reliance on operational patterns and promote deep understanding of mathematical equivalence through four key components, outlined below, that have independently been shown to be effective. Multiple pilot studies have since found that the ICUE treatment intervention is successful in improving student understanding of mathematical equivalence (Byrd et al., 2015; Johannes et al. 2017; Johannes and Davenport, 2019).

1. Nontraditional arithmetic practice (McNeil, Fyfe, & Dunwiddle, 2015; McNeil et al., 2011), in which operations are shown on the right side of the equal sign (e.g., __ = 3 + 6);
2. Lessons that first introduce the equal sign outside of arithmetic contexts (e.g., "28 = 28"; 2 apples = 2 apples) before introducing arithmetic expressions with operators (e.g., Baroody & Ginsburg, 1983),
3. Concreteness fading exercises in which concrete, real-world, relational contexts (e.g., sharing stickers, balancing a scale) are gradually faded into the corresponding abstract mathematical symbols (e.g., Fyfe, McNeil, Son, & Goldstone, 2014); and
4. Activities that require students to compare and explain different problem formats and problem-solving strategies (e.g., Carpenter, Franke, & Levi, L. 2003).

The current study. Previous findings (Byrd et al., 2015; Johannes et al. 2017; Johannes and Davenport, 2019) support greater learning outcomes for students who participate in the Treatment intervention activities. In the current study, we look at the learning trajectories for students in each condition. We examine not only whether and how students improve over the course of the intervention activities, but also how students with different initial levels of proficiency progress throughout the activities. We used matched pre- and post-intervention assessments to identify three categories of pre- and post-intervention performance.

Low-level performance: Students evidenced low-level performance on pre- or post-intervention assessments if they gave correct responses on 4/14 or fewer items (i.e., below 33% accuracy).

Mid-level performance: Students evidenced mid-level performance on pre- or post-intervention assessments if they gave correct responses on between 5/14 to 9/14 items (i.e., 33-66% accuracy).

High-level performance: Students evidenced high-level performance on pre- or post-intervention assessments if they gave correct responses on 10/14 or more items (i.e., above 66% accuracy).

Based on these divisions, we explore the following questions:

1. Are students assessed as low performing, pre-intervention, in the Treatment condition more likely to become high performing, post-intervention, than those in the Active Control condition?
2. Do students that improve from low to high performing, independent of condition, demonstrate higher accuracy on transfer measures of mathematical knowledge?
3. Can we predict successful learning at intermediate points in the intervention?

Assessing students' knowledge of equivalence. We assessed second-grade students' knowledge of mathematical equivalence before and after the intervention training with researcher-developed measures of equation encoding, equation solving, and defining the "=" symbol used in previous work by McNeil and colleagues (Johannes et al., 2017, 2019; McNeil et al., 2012; McNeil & Alibali, 2005b).

Equation encoding. The encoding measure consisted of recalling four math expressions (e.g., 2 + 6 =2 + _) presented one at a time. Each expression was visible for five seconds and students were instructed to remember and write down exactly what they saw. Responses were coded as correct if the student wrote the equation exactly as shown (i.e., the correct numbers and symbols in the correct order).

Equation solving. The equation solving measure consisted of eight equations with operations on both sides of the equal sign (e.g., 3 + 5 + 6 = 3 + __). For a response to be coded as correct, a student needed to write the value that would make the equivalence relation hold.

Defining = symbol. The defining the equal sign measure prompted students to write responses to three questions about the equal sign symbol (=): 1) What is the name of this math symbol? 2) What does this math symbol mean? And, 3) Can it mean anything else? Teachers read each question aloud and waited for students to write their responses before moving on to the next question. Responses were coded as correct (i.e., relational) definitions if the response defined the equal sign as relating two sides of the equation (e.g., two amounts are the same, something is equivalent to another thing).

Methods

Design

We used a randomized control trial design to examine the impacts of the ICUE intervention training relative to an active control program and analyzed learning outcomes at the student level to assess individual learning trajectories. Teachers were randomly assigned to use the either the ICUE intervention or Active Control materials. The Active Control consisted of workbook activities to control for time on task and contained non-traditional arithmetic practice but not the additional components present in the Treatment ICUE condition, described above.

Participants. Eighty-seven second-grade teachers (45 Treatment, 42 Active Control) used the activities in their classrooms in California. Class sizes ranged from 18 to 28, and we analyzed data from 922 students who completed the ICUE activities and 887 students who completed the Active Control activities and measures. Additionally, we collected student workbooks from 330 Treatment and 297 Active Control students in the sample – approximately 8 students per classroom – in order to examine students' engagement with the intervention materials at intermediate timepoints.

Procedures and Materials

The procedures for ICUE Treatment and Active Control conditions were identical, differing only in the content of the materials used by teachers and students. Each teacher received training on the study purpose, features of the activities, and strategies for integrating the activities into their typical mathematics curriculum.

Prior to starting the study, participating teachers completed online surveys assessing their mathematics teaching experience and classroom structure and dynamics.

After administering a pre-test, teachers used the study materials for approximately 15 minutes twice each week for 16 weeks. In both conditions, teachers were asked to use the study materials to supplement, rather than replace current math instruction, and to limit the duration of the activities to 20 minutes per session. Weekly implementation logs, completed by teachers, and samples of student workbooks confirm that all teachers in our sample administered every session within the general parameters laid out by the study instructions.

Proximal measures of mathematical equivalence. After completing the 32 sessions, teachers administered the same pre-intervention measure of mathematical equivalence understanding, which included the equation encoding and solving items reported here, along with an item prompting children to name and define the "=" symbol. We administered an additional researcher-developed measure of transfer word problems as well as a measure of computation fluency, but do not report these here.

Distal transfer to mathematical explanations. We selected two MARS items that tested second-grade students' understanding of mathematical equivalence, described below. Items were scored by project staff following scorer training, calibration, and reliability procedures established by MARS (Foster & Noyce, 2004).

Incredible Equations. In this task, students are asked to fill in the missing parts of equations such as "__ + 8 + __ = 16" and "11 + 5 = __ + 8." Students are asked to explain how they know their answer is correct. When 6,305 students took the task in 2007, the mean score was 6.08 out of 10 with a standard deviation of 2.5 (MARS, 2007).

Agree or Disagree? In this task, students are asked if they agree or disagree with two number sentences: "8 + 5 = 5 + 8" and "6 – 4 = 4 – 6". Students are asked to explain their answers using words, numbers, or pictures. MARS administered this task to 4,585 second graders in 2004 and found the mean score was 3.10 out of 6 with a standard deviation of 1.94 (MARS, 2004).

Active Control. Teachers in the Active Control condition received a set of student workbooks and a teacher guide. A sample workbook page is presented in Figure 1a.

ICUE. Teachers in the ICUE Treatment condition received a set of student workbooks (sample page in Figure 1a), a teacher guide, a set of classroom manipulatives including balance scales and flashcards with accompanying interactive stickers (Figure1b).

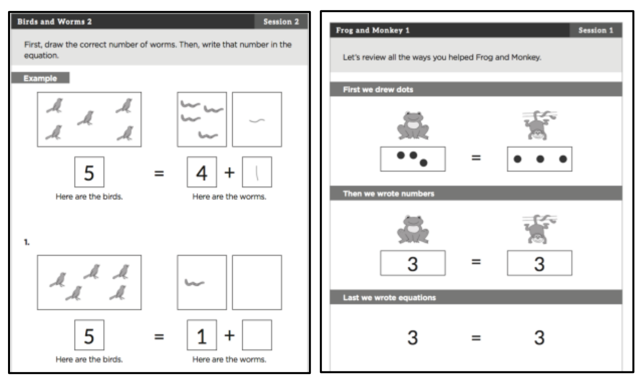

Figure 1a. Sample workbook page from the Active Control (left) and ICUE Treatment (right) condition materials.

Figure 1b. Materials, including workbooks and manipulatives, from the ICUE Treatment condition.

Results

Are low performing students at pretest in the Treatment condition more likely to become high performing at posttest than those in the Active Control condition?

Yes. Matched pre- and post-intervention assessments of mathematical equivalence confirm that students that performed "low" at pretest in the Treatment condition were more likely to perform "high" at posttest as compared to similar students in the Control condition. Pre-intervention, Active Control (Figure 2a) and Treatment (Figure 2b) samples show nearly identical proportions of low- (<33% accuracy), mid- (33-66% accuracy), and high-performing (>66% accuracy) students. However, a greater proportion of Treatment students were reclassified as high performing on the post-intervention assessment.

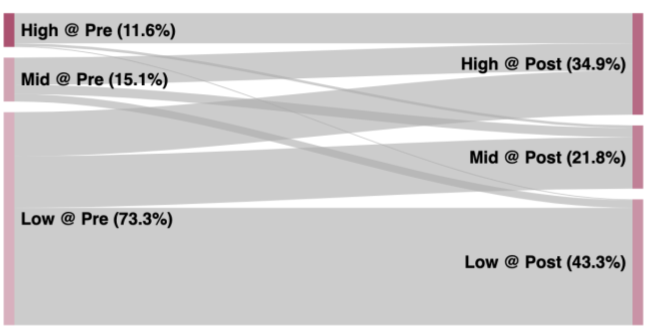

Figure 2a. Performance trajectories for students in the Active Control sample, based on high-, mid- and low-level performance on pre- and post-intervention assessments.

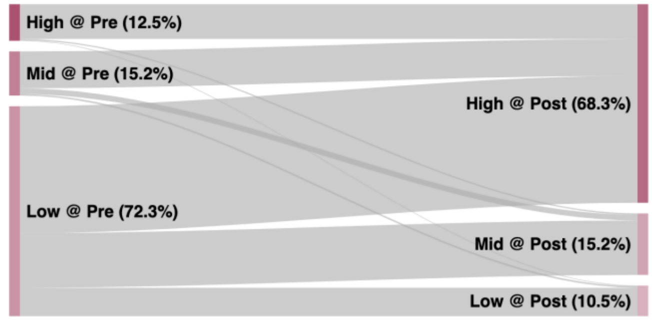

Figure 2b. Performance trajectories for students in the Treatment sample, based on high-, mid- and low-level performance on pre- and post-intervention assessments.

We used these learning trajectories to label categories of students. Our primary category of interest was termed "Successful Learners". These students showed low- or mid-level performance on the pre-intervention assessment and evidenced 25% or greater improvement, which typically resulted in high-level performance, on the post-intervention assessment. We defined three additional categories: "High achievers" showed high-level performance at both pre- and post-assessment; "Low achievers" showed low-level performance at pre- and post-assessment; and students in the "Other" category either showed a decrement in performance (45 students) or showed no change in their mid-level performance from pre- to post-assessment (46 students). Table 1 gives a breakdown of each intervention sample according to these categories.

Table 1. Percentage (and raw N) breakdown of Active Control and Treatment samples into trajectories of learners based on matched pre- and post-intervention assessments.

Category	Active Control	Treatment
High achievers	10.4% (92)	11.9% (110)
Low achievers	40.4% (358)	9.8% (90)
Learners	**42.3% (375)**	**75.2% (693)**
Other	7.0% (62)	3.1% (29)

A greater proportion of students in the Treatment condition were classified as Learners from pre- to post-intervention than in the Active Control condition. Conversely, a greater proportion of students in the Active Control condition were classified as Low achievers than in the Treatment condition (*Fischer's exact p* <.0001). The two conditions did not reliably differ in the proportion of High achiever students.

Is being a successful learner, independent of condition, predictive of higher accuracy on transfer measures of mathematical knowledge?

Yes. Our exploratory analysis demonstrated that students classified as Learners, independent of their assigned condition, were more likely to solve and explain transfer problems correctly, compared to Low achiever students and, in some cases, unclassified ("Other") students.

The *Incredible Equations* and *Agree or Disagree?* items from the MARS assessment series (MARS, 2004) served as our distal measures of knowledge transfer. Both items require students to solve a series of mathematics problems and to explain their solution or reasoning. We chose to examine students' overall scores for each item as well as isolated performance on the explanation sections for each item.

We compared differences in students' performance on these transfer assessments using a series of linear regression models to separately predict differences in total scores and explanation sub-scores for the Incredible Equations (Figure 3) and Agree/Disagree? (Figure 4) items based on condition, learner category, and their potential interaction.

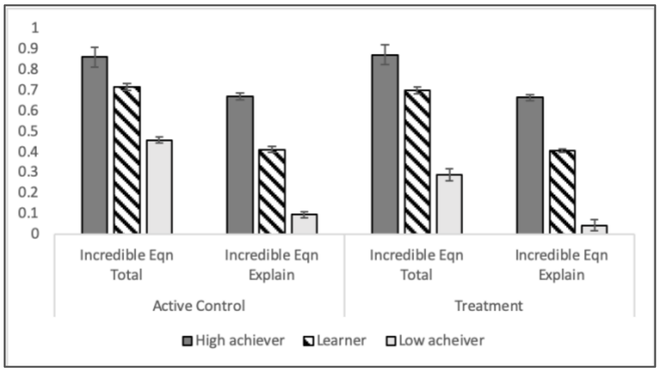

Figure 3. Performance on the "Incredible Equations" transfer item and explanation sub-score across condition and student learning classification.

Our regression models confirmed that students classified as Learners responded with higher accuracy compared to Low-achievers. For the Incredible Equations item, this relationship held true at the level of both the total score (β=0.18-0.41, SE=0.02-0.04, $p<.001$), and the explanation sub-score (β=0.23-0.57, SE =0.05-.07, $p<.001$). Condition, on its own, was not a reliable predictor of differences in either students' overall score or explanation sub-score for the Incredible Equations item but interacted with learner classification, such that the difference in overall score for Learners compared to Low achievers was greater in the Active Control condition (β=0.18, SE =0.04, $p<.001$). Finally, students classified as High achievers were predicted to score reliably higher, compared to learners, both overall (β=0.16, SE=0.02, $p<.001$) and on the explanation sub-score (β=0.28, SE=0.05, $p<.001$), and this predicted difference did not interact with condition.

Similarly, for the Agree/Disagree? item (Figure 4), students classified as Learners responded with higher accuracy compared to students classified as Low-achievers. This was the case for both the total score (β=0.35-0.48, SE =0.03-0.05, $p<.001$) and the explanation sub-score β =0.27-0.56, SE =0.04-0.06, $p<.001$). Assigned condition, on its own, was not predictive of differences in either the overall score or explanation sub-score. As before, High achievers scored reliably higher, compared to learners, both overall (β=0.23, SE=0.03, $p<.001$) and on the explanation sub-score (β=0.27, SE=0.04, $p<.001$), and this predicted difference did not interact with condition.

Thus, students who were classified as Learners based on matched pre- and post-intervention assessments were more likely to correctly solve and explain post-intervention transfer problems, compared to their Low achiever peers.

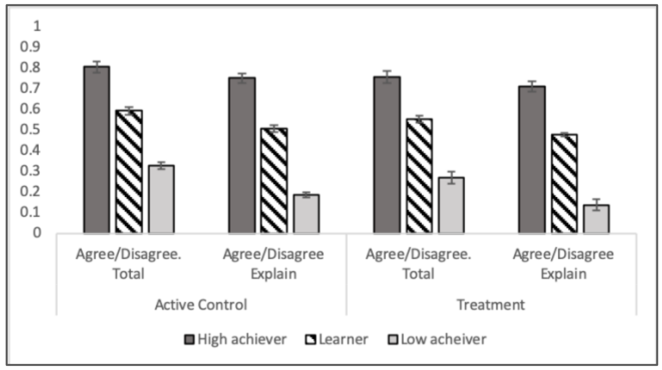

Figure 4. Performance on the "Agree/Disagree?" transfer item and explanation sub-score across condition and student learning classification.

Can we predict successful learning at intermediate points in the intervention?

After students completed all intervention activities and assessments, teachers were asked to collect and send in a sample of eight student workbooks from their classroom, stratified and randomly chosen by researchers based on pre-assessment scores. The workbooks were designed to be used during each of the 32 intervention sessions and, thus, a completed workbook reflected some aspect of engagement on the part of the student. Critically, workbooks in each condition featured different content, according to the design principles of the ICUE Treatment and Active Control interventions, discussed in the introduction. For each session, we coded every page of the workbooks according to a 3-point completion rating (complete, partially complete, incomplete/blank) that captured whether the student had engaged with the material for the session (i.e., irrelevant drawings or scribbles were not counted), but did not capture information about the accuracy of a student's workbook response.

We used this workbook completion data in a logistic regression model to predict whether a student would be classified as a successful Learner at the end of the intervention on the basis of assigned condition and his or her engagement with the intervention workbook activities. On this model, successful learner classification was reliably predicted by both total workbook completion: students who completed a greater proportion of the workbook pages were more likely to be classified as successful learners (β=3.41, SE=1.02, $p<.01$), and by assigned condition: students assigned to the Treatment condition were more likely to become learners (β=3.55, SE=1.16, $p<.001$), but not their interaction.

We were able to further predict successful Learners using a combination of workbook completion at the *halfway* point of the intervention (β=2.86, SE=0.95, $p<.01$) in combination with assigned condition (β=2.53, SE=1.22, $p<.05$). As before, the interaction between workbook completion and condition was not reliable. Thus, for both the Active Control and Treatment interventions, a student's engagement with the workbook activities, operationalized as workbook completion without monitoring accuracy, predicted whether they would show substantial learning gains from pre- to post-intervention assessment.

Conclusions

We tested two interventions designed to improve second-graders' understanding of mathematical equivalence. Although the randomly-assigned samples for each intervention condition started with equal proportions of low-performing students, our confirmatory analyses showed that those in the Treatment condition, which featured principle-based classroom and workbook activities, were more likely to become high-performing students at the end of the intervention, compared to those in the Active Control condition. Learners – students who progressed from low- to high-performance – in both intervention conditions were more likely to correctly solve and explain transfer problems, compared to low achiever students. While Treatment students were more likely to achieve Learner status, our exploratory analyses suggested that condition, on its own, was not predictive of a students' success on the transfer items. Consistent with this finding, previous work from McNeil and colleagues' (McNeil et al. 2011) reports that some students are able to generalize key concepts from exposure to non-traditional arithmetic practice – the primary content of the Active Control condition – alone. Thus, it is likely that Learners in the Active Control condition have similar depth of knowledge to Learners in the Treatment condition.

Additional exploratory analyses revealed that measures of engagement with workbook activities predicted whether a student would be classified as a Learner at the end of the intervention. This relationship held for workbook engagement measured as early as halfway through the intervention. This measure of intervention engagement doesn't reflect all the practice and exposure that Treatment students also engaged through in-class activities, or the accuracy with which students responded to workbook activities and practice problems, but the predictive relationship between attempting workbook activities and learning trajectory is nevertheless informative.

Identifying early predictors of successful learning or, conversely, potential warning signs by monitoring student engagement with materials could serve to alert teachers to the need for feedback (cf. McNeil & Alibali, 2000) and may be especially feasibly within online learning systems. In future work, we plan to address the time course of student engagement and its relationship to successful learning in these types of targeted interventions.

Finally, the current study focused on two subpopulations of students – successful learners in contrast to low achievers – and effectively ignored the small but well-defined subsample of high achieving students. These high achievers were already approaching ceiling on pre-intervention assessments and continued to excel on post-intervention assessments. This subgroup also demonstrated the highest performance on distal transfer measures, independent of condition.

This work contributes to a growing body of findings on the ICUE Treatment and Active Control interventions, including confirmatory efficacy findings in authentic classroom settings (Davenport et al., under review), and exploratory analyses examining the relationship between students' equation encoding and solving abilities before and after participating in each intervention program. Follow-up analyses will explore the depth of learning achieved by this already high-performing group in order to determine whether they are beyond the scope of the intervention materials or have succeeded in learning deeper concepts from the activities. In future iterations of work on the ICUE Treatment intervention activities, we aim to refine and adapt these highly efficacious materials for students and teachers to use online.

Acknowledgments

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.

References

Alibali, M. W., Phillips, K. M., & Fischer, A. D. (2009). Learning new problem-solving strategies leads to changes in problem representation. *Cognitive Development, 24*, 89-101.

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

Behr, M., Erlwanger, S., & Nichols, E. (1980). How children view the equal sign. *Mathematics Teaching, 92*, 13-15.

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, *2*, 123-152.

Byrd, C. E., McNeil, N. M., et al. (2015). Pilot Test of a Comprehensive Intervention to Improve Children's Understanding of Math Equivalence. AERA, Chicago, IL.

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

Charles, R. I. (2005). Big ideas and understandings as the foundation for elementary and middle school mathematics. *NCSM: J. of Mathematics Education Leadership*, 8(2), 9-24.

Davenport, J., Kao, Y., Johannes, K., Hornburg, C.B., & McNeil, N.M. (under review). Improving Children's Understanding of Mathematical Equivalence: An Efficacy Study.

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

Foster, D. & Noyce, P. (2004, January). The Mathematics Assessment Collaborative: Performance testing to improve instruction. *Phi Delta Kappa*, 367-374.

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*, 9-25.

Johannes, K., Davenport, J., Kao, Y, Hornburg, C.B., & McNeil, N.M. (2017). Promoting children's relational understanding of equivalence. *Proc. of the 39th Annual Meeting of the Cognitive Science Society*, 600-605.

Johannes, K., & Davenport, J. (2019). Targeted mathematical equivalence training lessens the effects of early misconceptions on equation encoding and solving. *Proc. of the 40th Annual Meeting of the Cognitive Science Society*, 492-499.

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*, 166-176.

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

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.

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

McNeil, N. M. (2014). A "change-resistance" account of children's difficulties understanding mathematical equivalence. *Child Development Perspectives*, *8*, 42-47.

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.

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*, 451-466.

McNeil, N. M., & Alibali, M. W. (2005). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development*, *76*, 883-899.

McNeil, N. M., Chesney, D. L., et al. (2012). Organizing addition knowledge around equivalent values facilitates understanding of math. equivalence. *J. of Educational Psychology*. *104*, 1109-1121.

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*, 1620-1633.

McNeil, N. M., Fyfe, E. R., & Dunwiddie, A. E. (2015). Arithmetic can be modified to promote understanding of mathematical equivalence. *J. of Educational Psychology*, *107*, 423-436.

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

Seo, K.-H., & Ginsburg, H. P. (2003). "You've got to carefully read the math sentence...": Classroom context and children's interpretations of the equals sign. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 161-187). Mahwah, NJ: Erlbaum.