Abstract Title Page

Not included in page count.

Title: Understanding the Equals Sign as a Gateway to Algebraic Thinking

Author(s):

Percival G. Matthews, Vanderbilt University; Bethany Rittle-Johnson, Ph.D., Vanderbilt University; University; and Roger S. Taylor, Ph.D., Vanderbilt University; Katherine L. McEldoon, Vanderbilt University

Abstract Body

Limit 5 pages single spaced.

Background/context:

Description of prior research, its intellectual context and its policy context.

Mathematical equivalence is a foundational concept of algebraic thinking that serves as a key link between arithmetic and algebra (MacGregor & Stacey, 1997). Typically represented by the '=' symbol, equivalence is the principle that two sides of an equation represent the same value. True understanding of equivalence requires thinking about the relation between the two entities on either side of the equal sign (i.e., relational thinking). Several studies have shown that knowledge of the concept supports greater algebraic competence, including equation-solving skills and algebraic reasoning (Kieran, 1992; Knuth, Stephens, McNeil & Alibali, 2006; Steinberg, Sleeman & Ktorza, 1990).

Unfortunately, numerous past studies have also pointed to the difficulties that elementary-school children have understanding equivalence. Although elementary school students have some basic understanding of what it means for quantities to be equal, these children often interpret the equals sign as simply an operator signal that means ''adds up to'' or ''gets the answer.'' (Baroody & Ginsburg, 1983; Kieran, 1981; Rittle-Johnson & Alibali, 1999; Sfard & Linchevski, 1994). This *operational* view of the equal sign can impede the development of a *relational* view of the equal sign. An operational view of the equal sign often persists for many years, and students who have this view often have difficulty solving equations (Knuth, et al., 2006). As a result, most elementary-school children reject equations not in a standard "a + b = c" structure as false (e.g., 3 = 3 and 3 + 5 = 5 + 3). A long list of studies spanning the last 35 years of research has shown that a majority of first through sixth graders treated the equal sign operationally when solving equations not in a standard "a + b = c" structure, often leading to both computational and conceptual errors (Alibali, 1999; Behr, Erlwanger, & Nichols, 1980; Falkner, Levi, & Carpenter, 1999; Jacobs, et al., 2007; Li, Ding, Capraro, & Capraro, 2008; McNeil, 2007; Perry, 1991; Rittle-Johnson, 2006; Rittle-Johnson & Alibali, 1999; Weaver, 1973).

Despite the importance of the topic and the years of research dedicated to its study, few researchers have used psychometrically validated measures when investigating mathematical equivalence. Indeed, this measurement problem is prevalent in math education more generally – for example, Hill & Shih (2009) found that less than 20% of studies published in the Journal for Research in Mathematics Education over the past 10 years had reported on the validity of the measures. The lack of valid measures makes it difficult to evaluate changes in knowledge over time or the effectiveness of interventions. Cognizant of these facts, we have been developing an instrument for measuring school students' knowledge of mathematical equivalence using the assessment development framework laid out by the AERA/APA/NCME *Standards for Educational and Psychological Testing.*

Purpose / objective / research question / focus of study:

Description of what the research focused on and why.

In this study, we wanted to examine whether success on items testing basic equivalence knowledge, such as the meaning of the equal sign and ability to solve problems such as 3 + 5 = 4

+ _, predicted success on items testing more advanced algebraic thinking (i.e. principles of equality and solving equations that use letter variables). This investigation is a follow up study to our initial efforts to design an instrument to measure children's understanding of equivalence (Rittle-Johnson, et. al. under review). This replication and extension with a new sample also provides evidence for the validity and generalizability of our instrument.

We had two specific predictions about the relations between basic-level and advanced-level knowledge items. First, we expected that the relative difficulty of the two types of knowledge would be born out on the Rasch model. That is, we expected that our empirically derived difficulty scores would be higher for the advanced-level items than for the basic-level item. Second, we expected that performance on basic-level items could be used to predict performance on advanced-level items.

Setting:
Description of where the research took place.

Data was collected during class time in 13 second- through sixth grade classrooms in two suburban, public schools in Tennessee. Data were collected at a single time point for each class.

Population / Participants / Subjects:
Description of participants in the study: who (or what) how many, key features (or characteristics).

224 second- through sixth-grade students participated near the end of the school year. Of the students who completed the assessment, 53 were in second grade (23 girls), 46 were in third grade (25 girls), 29 were in fourth grade (14 girls), 59 were in fifth grade (26 girls), and 37 were in sixth grade (16 girls). The mean age was 10.2 years (SD = 1.6; Min = 7.7; Max. = 14.1). The students were predominantly Caucasian; approximately 2% of students were from minority groups. The schools served a working- to middle-class population.

Intervention / Program / Practice:
Description of the intervention, program or practice, including details of administration and duration.

This study focused on instrument development, so there was no intervention. Thus, what follows describes the creation and administration of our assessment.

In a previous study, we followed the construct modeling approach of Wilson (2005), using item response theory (IRT) and to create a criterion-referenced framework for determining students' understanding of mathematical equivalence. Specifically, we previously 1) developed a construct map covering students' knowledge of mathematical equivalence (Table 1), 2) used the construct map to develop a comprehensive assessment, 3) administered the assessment to students in Grades 2 to 6, and then 4) used the data to evaluate the construct map and the assessment (Rittle-Johnson et. al., under review).

We developed two comparable forms of an assessment tool from a pool of assessment items selected from past research, state and national assessments, and standardized tests. The items took an assortment of formats, including multiple choice, fill in the blank, and short answer. Both forms of the assessment were comprised of three sections, based on the three most commonly used types of items in the literature:

- *Equal-sign items* – These items were designed to probe students' explicit knowledge of the equals sign as an indicator of equivalence.

- *Equation-structure items* – These items were designed to probe students' knowledge of valid equation structures.
- *Equation-solving items* – These items were designed to probe students' abilities to solve equations.

In the construct map, we proposed four levels of increasing knowledge guided in part by the benchmarks proposed by Carpenter, Franke & Levi (2003). We generated items to cover each of the following four levels in order of increasing difficulty:

1. *Rigid operational*, in which children hold an operational view and can only solve problems in the standard "a + b = c" format;
2. *Flexible operational,* in which children hold an operational view, but can solve equations in some nonstandard formats (e.g. c = a + b)
3. *Relational with computational support*, in which a nascent relational coexists with an operational view, allowing students to solve equations with operations on both sides (e.g. a + b + c = a + _)
4. *Relational without need to compute* (full relational), in which a relational view predominates and children demonstrate understanding for the arithmetic properties of equivalence.

We defined *advanced-level* items as Level 4 items that test students' understandings of the principles of equality and their abilities to solve equations that use letter variables. We defined *basic-level* items as those falling on Levels 1-3, with the exception of one Level 3 item that used a letter variable.

Based on feedback from a panel of experts in mathematics education and empirical evidence of item performance, we made some minor changes to the original assessments we designed for Rittle-Johnson, Taylor, Matthews & McEldoon (under review). The most significant change was the addition of several more advanced-level items. For example, we added a new section of items that focused on students' understanding of the principles of equivalence. For instance, one such problem began by stating "25+14=39 is true." It then asked, "Is 25+14+7=39+7 true or false?" Students were asked to circle either "True," "False," or, "Don't Know," and to explain the answers that they chose. These problems were designed to test students' knowledge of the arithmetic properties of equivalence, which hold that an equivalence relationship remains true as long as an identical operation is performed on both sides of the equal sign. These types of problems have been cited as addressing the types of thought that underlie formal transformational algebra (Kilpatrick, Swafford, & Findell, 2001).

A second set of additional problems addressed the principles of equivalence using letter variables (literals) that are typically seen in formal algebra. For instance, one asked, "Find the value of *c*," for the equation $c + c + 4 = 16$. These items are important because the use of variables—particularly multiple instances of the variable— tests whether students comprehend that a variable represents a specific and constant number value.

The final versions of the assessments each consisted of 39 items, 9 of which qualified as advanced-level items and 22 of which qualified as basic-level items. The total does not sum to39, because six of the remaining items were Level 4 items that neither explicitly tested the arithmetic principles of equality nor used letter variables, and one item was a Level 3 item that used a letter variable. The assessments were administered on a whole-class basis by a member of the project team. Completion of the assessment required approximately 45 minutes and was performed within a single class period. Test directions were read aloud for each type of item in

2nd grade classrooms to minimize the possibility that reading level would affect performance. Otherwise, test administration was identical across grade levels.

Research Design:
Description of research design (e.g., qualitative case study, quasi-experimental design, secondary analysis, analytic essay, randomized field trial).

This study focused on measurement development and utilized item response theory (IRT) and the construct modeling approach of Wilson (2005) to create a criterion-referenced framework for determining students' understanding of mathematical equivalence. Specifically, we previously 1) developed a construct map covering students' knowledge of mathematical equivalence (Table 1), 2) used the construct map to develop a comprehensive assessment, 3) administered the assessment to students in Grades 2 to 6, and then 4) used the data to evaluate the construct map and the assessment (Rittle-Johnson et. al., under review).

In the construct map, we proposed four levels of increasing knowledge guided in part by the benchmarks proposed by Carpenter, Franke & Levi (2003). We used factor analysis to confirm the unidimensionality of our construct and Rasch analysis to ensure that item difficulty levels operated as hypothesized. Our analyses suggested that we had developed a very promising assessment of equivalence knowledge. We made minor adjustments to the construct map and assessment based on this first round of data, and this study was carried out to further test the assessment with a different sample of students.

Data Collection and Analysis:
Description of the methods for collecting and analyzing data.

After test administration, all items on the assessment were coded as binary responses (1 = correct). Next, we assessed the degree of internal consistency among the items using Cronbach's alpha (α's > .94). We supplemented this measure of internal consistency with confirmatory factor analysis to assess the dimensionality of the constructed measured. Then, we fit the data to a Rasch model. This model is a member of the IRT family of analytic models and simultaneously plots difficulty levels and student skill levels on a logit scale. This allows us to calculate the probability that a participant will get a given question right given his/her ability level. Of a total of 43 test items, four were dropped because multiple fit indicators suggested that they failed to have good psychometric properties. For the remaining 39 items, we performed a univariate ANCOVA to investigate the relations between basic level knowledge and advanced knowledge.

Findings / Results:
Description of main findings with specific details.

As detailed above, we tested two primary hypotheses: 1) that the hypothesized relative difficulty of the two types of knowledge would be born out empirically; and 2) that performance on basic-level items could be used to predict performance on advanced-level items.

Validity of Relative Difficulties. An item-respondent map (i.e., a Wright Map, see Figure 1) generated by the Rasch model was used to evaluate our construct map. Our Wright Map places *respondents or participants* on the left side of the vertical axis and place test items on the right side of the axis. Participants of higher ability are located the upper portion of the map, while those of lesser ability are located on the lower portion. Similarly, on the right, *items* of greater difficulty are located near the top of the map and those of lesser difficulty are lower on the map. The locations of participants and respondents are measured in logits (i.e., log-odds units), which for a given item-participant pairing is calculated as the natural logarithm of the participant's estimated probability of success divided by the estimated probability of failure on an item.

Our hypotheses about the relative difficulties of the various items were largely borne out by the empirical data – the items we classified as higher level were place higher on the scale than more basic level items, with a few exceptions. These exceptions gave us potentially valuable feedback for reassessing the difficulty of some of the items.

As in prior studies, many students failed to demonstrate a relational understanding of equivalence. On average, participants were only 27 percent accurate on Level 4 items across grade and only 57 percent accurate on level 3 items as compared to 76, and 85 percent accurate on Level 2 and Level1, items respectively.

Predictive relations between advanced and basic-level items. To investigate the correlation between proficiency with basic-level items with higher-level items, we ran a univariate ANCOVA, with performance on higher level problems as a dependent measure and performance on lower level items and grade as predictor variables.
Performance on lower level items was significantly associated with success on higher level items, $F(1, 217) = 102.40$, $p < .01$, $\eta^2 = .32$. There was also a significant effect for grade, $F(4, 217) = 6.46$, $p < .01$, $\eta^2 = .11$. Thus, performance on lower level items was able to predict 32 percent of the variance in performance on higher-level items. Of course, we found this relationship using cross-sectional data using measures that were administered within a single sitting. Thus, we interpret our results with caution and recognize the need to replicate our results on future data sets to confirm our findings. On the whole, the data replicate and extend our previous investigation using a new sample of participants.

Conclusions:
Description of conclusions and recommendations based on findings and overall study.

Previous studies have suggested that students' understandings of the equals sign should inform their use of algebraic strategies (Knuth et. al., 2006; Alibali, Knuth, Hattkudur, McNeil & Stephens, 2007). We sought corroborating evidence for this claim by examining whether understanding of basic-level difficulty items that tap equivalence knowledge could also predict facility with more difficult higher-order items in elementary school. Our study had the added virtue of using a psychometrically sound measurement instrument. We found that proficiency with basic-level items explained much of the variance in student success on more advance problems. Our findings provides evidence for the assumption that supporting basic algebraic thinking in elementary school may improve more traditional algebraic competence in middle school.

Our instrument has the potential to be of considerable benefit to educational researchers and practitioners. First, it can provide a valid metric for the evaluation of experimental interventions. The development of such methods is integral to SREE's mission to "increase the capacity to design and conduct investigations that have a strong base for causal inference". Second, it can provide a tool for teachers to identify students' knowledge levels. Research has shown a) that teachers often over-estimate what their students know about the equal sign (Falkner, et al., 1999); b) that raising teachers' awareness of their students' lack of understanding can be a powerful motivator for changing their instruction (Carpenter, Fennema, Peterson, Chiang, & Loef, 1989; Fennema, Carpenter, Franke, Levi, & et al., 1996; Jacobs, et al., 2007); and c) that differentiated instruction has been shown to improve student achievement (e.g., Mastropieri, et al., 2006; Richards & Omdal, 2007), but that teachers often lack the tools for identifying students' knowledge levels and customizing their instruction (e.g., Houtveen & Van de Grift, 2001). Our instrument can help address each of these issues.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology, 35*(1), 127-145.

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 equals sign. *Mathematics Teaching*(92), 13-15.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C. P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal, 26*(4), 499-531.

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

Hill, H. C. & Shih, J. (2009) Examining the Quality of Statistical Mathematics Education Research. *Journal for Research in Mathematics Education*, 40(3), 241-250.

Houtveen, T., & Van de Grift, W. (2001). Inclusion and adaptive instruction in elementary education. *Journal of Education for Students Placed at Risk, 6*(4), 389-409.

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., Franke, M. L., Levi, L., & et al. (1996). A longitudinal study of learning to use children's thinking in mathematics instruction. *Journal for Research in Mathematics Education, 27*(4), 403-434.

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.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics, 12*(3), 317-326.

Kieran, C. (1992). The learning and teaching of school algebra. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 390-419). New York: Simon & Schuster.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). Adding it up: Helping children learn mathematics. Washington, DC: National Academies Press.

Knuth, E. J., Stephens, A. C., McNeil, N. M., & Alibali, M. (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.

MacGregor, M., & Stacey, K. (1997). Students' understanding of algebraic notation. *Educational Studies in Mathematics, 33*(1), 1-19.

Mastropieri, M. A., Scruggs, T. E., Norland, J. J., Berkeley, S., McDuffie, K., Tornquist, E. H., et al. (2006). Differentiated curriculum enhancement in inclusive middle school science: Effects on classroom and high-stakes tests. *The Journal of Special Education, 40*(3), 130-137.

Matthews, P., & Rittle-Johnson, B. (2007). To teach by concept or by procedure? Making the most of self-explanations. In D. S. McNamara & G. Trafton (Eds.), *Proceedings of the 29th Meeting of the Cognitive Science Society*.

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.

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.

Perry, M. (1991). Learning and transfer: Instructional conditions and conceptual change. *Cognitive Development, 6*(4), 449-468.

RAND Mathematics Study Panel. (2003). *Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education*. Santa Monica, CA: RAND.

Richards, M. R. E., & Omdal, S. N. (2007). Effects of tiered instruction on academic performance in a secondary science course. *Journal of Advanced Academics, 18*(3), 424-453.

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

Rittle-Johnson, B. & Alibali, M.W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology*, *91*, 1-16.

Rittle-Johnson, B., Siegler, R.S. & Alibali, M.W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93, 346-362.*

Rittle-Johnson, B., Taylor, R.T., Matthews, P.G., & McEldoon, K. (under review). Mastering mathematical equivalence: A developmental model. *Journal of Educational Psychology.*

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* (pp.161–187). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Sfard, A. & Linchevski, L. (1994). The gains and pitfalls of reification – the case of algebra. *Educational Studies in Mathematics*, 26, 191–228.

Steinberg, R. M., Sleeman, D. H., & Ktorza, D. (1990). Algebra students' knowledge of equivalence equations. *Journal for Research in Mathematics Education, 22*, 112-121.

Weaver, J. F. (1973). The symmetric property of the equality relation and young children's ability to solve open addition and subtraction sentences. *Journal for Research in Mathematics Education, 4*(1), 45-56.

Wilson, M. (2005). *Constructing Measures: An Item Response Modeling Approach*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Appendix B. Tables and Figures
Not included in page count.

Table 1: Equal sign construct map

Levels	Performances	Items Expected to Get Correct	Response Exemplars
4. Relational without need to compute	Successful with equations with large numbers because can use relation between expressions, rather than computing. Understand principles of equivalence (doing same thing to both sides).	1) 67 + 84 = _ + 83 2) What is the best definition of the equal sign? 3) *Without subtracting* the 9, can you tell if the statement below is true or false? 76 + 45 = 121 is true. Is 76 + 45 – 9 = 121 – 9 true or false? True False Can't tell without subtracting How do you know?	1) 68 2) The equal sign means two amounts are the same 3) You did the same thing to both sides, so they're the same.
3. Relational with computational support	Successful with equations with operations on both sides, by computing solutions, and knows a relational definition of the equal sign, although it co-exists with an operational definition.	1) _ + 9 = 8 + 5 + 9 2) What does the equal sign mean? 3) For this example, decide if the number sentence is true. Then, <u>explain how you know.</u> 4 + 1 = 2 + 3 True False Don't Know How do you know?\	1) 13 2) "the same as", but may give second definition (e.g. the answer) 3) True You get five on each side, so they're the same.
2. Flexible Operational	Successful with equations with operations on the right (c = a + b) because they are just "backwards" but continues to think of equal sign	1) 7 = _ + 3 2) Is this statement true or false? 1 quarter = 25 pennies True False Don't Know 3) After each problem, circle True, False, or Don't Know.	1) 4 2) True 3) True

	operationally, or in other non-relational ways.	4 = 4 + 0 True False Don't Know	
1. Rigid Operational	Only successful on equations in standard "a + b = c" format and think of equal sign operationally (e.g., it means "get the answer").	1) 6 + 2 = _ 2)Which of these pairs of numbers is equal to 3 + 6? Circle your answer. 2+7 3+3 3+9 none 3) After each problem, circle True, False, or Don't Know. 7 + 6 = 0 True False Don't Know	1) 8 2) 2 + 7 3) False

Figure 1: Wright map of respondents and item difficulties

```
            PERSONS - MAP - ITEMS
                 <more>|<rare>
 5                     +
                       |
                       |
                  XX T |
                       |T
 4            XXXXXX   +
                       |
               XXXXX   |   ESDefineRelational_L4      STSameonBoth_L4
                       |
            XXXXXXXX   |   STAddBothExplain_L4        STSubtractBothExplain_L4
 3                     +   STMultiplyBothExp_L4
            XXXXXXXX   |
             XXXXXXX   |   OEVariableBothSides_L4     STRelatedExplain_L4
                     S |
             XXXXXXX   |
 2         XXXXXXXXX   +S  ESDefine_L3
          XXXXXXXXXX   |
        XXXXXXXXXXXX   |
           XXXXXXXXX   |   ESBestDef_L4               OEVariable_Multiple_L4
                           STSubtractBoth_L4
            XXXXXXXX   |
 1         XXXXXXXXX   +   STRelated_L4
                XXXX   |
               XXXXX   |   ESDefineMoney_L2           OERelatedLeft_L4
                           OERelatedRight_L4          STAddBoth_L4
            XXXXXXXX   |
               XXXXX M |   STBothExplain_L3
 0         XXXXXXXXX   +M  STMultiplyBoth_L4
                XXXX   |   OEBothLeft5_L3
                 XXX   |   OEBothRight5_L3            OEBothSub_L3
                           STBoth5_L3
                 XXX   |   OEBothRight_L3             OEOneVariable_L3
                           STIdentity_L2
               XXXXX   |   ESRateRelational_L3        OEBothLeft_L3
                           STCommutitive_L3
-1                XX   +   STAddSame_L4               STBoth_L3
                           STRightExplain_L2
                 XXX   |
                 XXX   |
          XXXXXXXXXX   |
                XXXX   |
-2            XXXXXX S+S   ESMoney_L2                 STZero_L2
         XXXXXXXXXXX   |
        XXXXXXXXXXXX   |   ESEqual_L1                 STRight_L2
                       |
        XXXXXXXXXXXX   |
-3                     +   OERight_L2
            XXXXXXXX   |   STFalse_L1
                 XXX   |
                       |
                       |
-4              XXXX   +   ESMoneyStatement_L2
                     T|T
                       |
                       |
                       |
-5                     +
```