Loehr, A. M., Rittle-Johnson, B., Star, J. R., & Desharnais, C. (2018, April). Developing a more comprehensive measure of formal algebra knowledge. Poster presented at the American Educational Research Association (AERA), New York City, NY.

Title: Developing a More Comprehensive Measure of Formal Algebra Knowledge

Authors: Abbey M. Loehr, Bethany Rittle-Johnson, Jon R. Star, & Chad Desharnais

Abstract (120 word limit)

Proficiency in algebra is critical to academic, economic, and life success. Competency in algebra requires developing conceptual knowledge and procedural flexibility. However, existing assessments rarely assess these two types of knowledge. We examined the psychometric properties of a new, comprehensive Algebra I assessment with conceptual knowledge, procedural knowledge, and procedural flexibility subtests. Seventy-six Algebra I students completed the 30-item assessment at the end of the school year. Average accuracy was 49% correct. Internal consistency was very good ($\alpha = .85$). Accuracy was significantly correlated with course grades, $r = .51$, as well as state test scores, $\rho = .57$. Valid and reliable measures of algebra knowledge are critical for evaluating interventions and developing theories of algebra learning.

Keywords:

Word count (2,000 word limit):

Developing a More Comprehensive Measure of Formal Algebra Knowledge

Objectives

Proficiency in algebra is critical to academic, economic, and life success (Adelman, 2006; NMAP, 2008). Unfortunately, national and international assessments have drawn attention to pervasive student difficulties in algebra (Schmidt et al., 1999). Because of the importance of proficiency with algebra, algebra learning has been an area of scholarship for many years. However, very few comprehensive measures of formal algebra knowledge exist for researchers to test the effectiveness of interventions. Further, the reliability and validity of measures is often not examined.

The validity of inferences that can be drawn from studies without information about the reliability and validity of key outcomes is limited because it is unknown whether these outcomes measure what they claim to (Hill & Shih, 2009). Thus, the development of valid and reliable measures of formal algebra knowledge is critical for evaluating the impact of interventions and developing theories of algebra learning. The goal of the study was to construct and examine the psychometric properties of a new, comprehensive Algebra I assessment. Given the lack of existing assessments, this important work is among the first to provide evidence for a reliable and valid measure, a process heavily emphasized by the AERA/APA/NCME (1999) testing standards.

Theoretical Framework

Competency in algebra requires developing both conceptual (i.e., knowledge of abstract concepts and general principles) and procedural knowledge (i.e., knowledge of mathematical strategies). Recent theories of algebra learning have focused on developing conceptual knowledge, building on early work by Kieran (1992). Conceptual knowledge is of critical importance; understanding of key concepts such as equivalence and variable is essential to success in algebra (Knuth, Stephens, McNeil, & Alibali, 2006).

Complementing knowledge of key concepts, success in algebra requires flexibility in the use of symbolic strategies. However, existing theories of algebra learning place less emphasis on symbolic strategy use. Working with symbolic strategies is essential in algebra learning. In particular, students need to develop procedural flexibility - knowing multiple strategies for solving a problem and selecting the most appropriate strategy for a given problem – as well as understand the conceptual rationale behind commonly used strategies. Learners who develop procedural flexibility are more likely to use or adapt existing strategies when faced with unfamiliar problems and to have greater conceptual knowledge (Blöte, Van der Burg, & Klein, 2001; Hiebert et al., 1996). Furthermore, procedural flexibility is a salient characteristic of experts in mathematics (Dowker, 1992; Star & Newton, 2009). Procedural flexibility is distinct from, but related to, conceptual and procedural knowledge of algebra (Schneider, Rittle-Johnson, & Star, 2011). For example, in a study on middle-school students' knowledge of solving linear equations, the latent variables for each knowledge type were correlated .63-.66, and model comparisons confirmed that the data was better represented using three distinct knowledge types rather than a single knowledge type. Further, conceptual and procedural knowledge of equation solving at the beginning of the unit contributed independently to developing procedural flexibility for equation solving by the end of the unit.

Evidence and theory suggests procedural flexibility and conceptual and procedural knowledge are all important for algebra learning. However, current measures of algebra knowledge rarely focus on conceptual knowledge or procedural flexibility. One researcher-based effort in developing a conceptual knowledge assessment for concept of variable did not achieve acceptable reliability (Genareo et al., 2016). Another did demonstrate some evidence of internal consistency for both conceptual knowledge and procedural flexibility, but focused on only a subset of Algebra I topics (Star et al., 2015). We developed a new, more comprehensive Algebra I assessment designed to assess procedural flexibility, conceptual knowledge and procedural knowledge. Our assessment broadly covers core Algebra 1 content, including linear equations, systems of equations, polynomials and factoring, and quadratic equations. In this paper, we describe the measure and report on the psychometric properties of the measure.

Method and Data Sources

Students from 4 Algebra I classrooms at a suburban public school in Massachusetts participated (N = 76 students; *M* age = 14 yrs; 50% female; 16% ethnic minorities, 26% were in Grade 8, considered by the district to be in the advanced math class, and 74% were in Grade 9, considered to be in the regular math class). No students had diagnosed learning disabilities, and all students were fluent in English. Their teachers had spent some time supporting procedural flexibility and conceptual knowledge through the comparison of multiple strategies. Teachers administered the assessment during a single class period at the end of their year-long Algebra 1 course.

Assessment. The assessment included 27 multiple-choice items and 3 short constructed-response items on core algebra topics (i.e., linear equations, systems of linear equations, polynomials and factoring, and quadratic equations). Items were modified from existing assessments and sampled to be representative of the most commonly assessed topics. All items were scored as correct or incorrect. The 3 short constructed response items were all procedural knowledge items and requested a numeric answer, which were scored as correct if the student provided the correct answer.

The assessment included three item types. Conceptual knowledge items (n = 10) targeted core concepts such as equivalence, linearity, variable, and solutions to systems of equations and quadratic equations. Procedural knowledge items (n = 10) required students to solve and graph linear equations, solve systems of equations, factor polynomials, and solve quadratic equations. Procedural flexibility items (n = 10) primarily focused on identifying the most efficient strategy (e.g., “On a timed test, which would be the BEST way to solve the problem below?”). Two of the items focused on knowledge of multiple strategies (e.g., “Which of the following would be mathematically okay way(s) to start solving the problem?”). See Figure 1 for sample items of teach type.

Criterion measures. To establish evidence of validity, we gathered criterion measures to examine how student scores on our algebra assessment were related to their scores on other mathematics assessments. Two criterion measures were gathered from school records: math course grade for the first semester of Algebra I and students’ performance category on the state test (i.e., needs improvement, proficient, or advanced on the MCAS) in the previous year.

Results

Overall assessment. On average, students solved 49% (SD = 20%) of problems correctly. The easiest items focused on connecting linear graphs to symbolic equations and properties of solutions to systems of equations. The hardest items required knowledge of solving quadratic equations and simplifying radicals in fraction form. Examples of easy and hard items from each item type along with performance data are shown in Figure 1. Students' scores on the procedural flexibility items (*M* = 42% correct) were similar to scores on the procedural knowledge items (*M* = 46% correct). Scores on the conceptual knowledge items (*M* = 60% correct) were higher than scores on both the procedural flexibility and procedural knowledge items, t (75) = 6.4 and 5.0, respectively, p's < .001.

The overall assessment is both reliable and valid (see Table 1). Internal consistency was very good (Cronbach's Alpha = .85). Criterion validity was examined by correlating accuracy on the assessment with students' course grades, state test scores, and course level (Advanced or Regular). Accuracy was significantly correlated with course grades, r (74) = .51, $p < .001$, as well as state test scores from the previous school year, Spearman's rho (71) = .57, $p < .001$. Students who had scored in the 'needs improvement' range (n = 2) on the state test solved on average 43% correct on the assessment, those in the 'proficient' range (n = 29) scored 36% correct, and those in the 'advanced' range (n = 42) scored 60% correct. Students in the advanced level of the course outperformed those in the regular level of the course (*M* = 76% vs. 40% correct, t (74) = 11.4, $p < .001$).

Although evidence for the reliability and validity of the overall assessment was strong, evidence for the reliability of individual item type sub-scores was weaker. For the procedural flexibility and conceptual knowledge sub-scores, internal consistency was acceptable ($\alpha = .66$); for procedural knowledge sub-scores, it was good ($\alpha = .79$). Students with higher procedural flexibility sub-scores also had higher conceptual and procedural knowledge sub-scores, r's (74) = .51 and .47, p's < .001, and conceptual and procedural knowledge sub-scores were correlated, r (74) = .59, providing some evidence for the validity of the individual item type sub-scores. The weaker internal consistency of the item type sub-scores reduces the strength of potential correlations with the item type sub-scores. See Table 1 for additional validity information.

Scholarly Significance

Proficiency in algebra is critical to the academic, economic, and life success of U.S. students (e.g., Adelman 2006). Students' scores on our comprehensive Algebra I assessment were modest after a year of algebra instruction, and conceptual knowledge scores were higher than procedural flexibility and procedural knowledge scores. The assessment was designed to tap strong knowledge of core Algebra I content, and results suggest that the students had not mastered much of the core content. The current sample was drawn from regular and advanced sections of Algebra I at a suburban public school in Massachusetts, almost all of whom had scored as proficient or advanced on the state test the year before. Given the difficulties low-performing students typically have with formal algebra, we would expect scores to be even lower with a more representative sample.

Our new, more comprehensive Algebra I assessment demonstrated strong reliability and good validity, similar to what was found in Star et al. (2015) on a less comprehensive measure. The validity of inferences that can be drawn from studies without information about the reliability and validity of key outcomes is limited because it is unknown whether these outcomes measure what they claim to (Hill & Shih, 2009). Thus, our more comprehensive measure of Algebra 1 content is critical for evaluating the impact of interventions and developing theories of algebra learning. Interventions to promote stronger formal algebra knowledge, including conceptual knowledge and procedural flexibility, are urgently needed.

Decomposing our assessment into sub-scores for procedural flexibility, conceptual knowledge and procedural knowledge is promising, but requires additional work. Given time constraints on the number of items that can be administered, it is difficult to administer a sufficient number of items of each type for the multiple topics covered in Algebra I. The reliability (internal consistency) of the sub-scores is acceptable for making group comparisons (e.g., comparing an experimental and control condition), but is problematic for examining individual differences (Thorndike & Thorndike-Christ, 2010). Others have had difficulty creating reliable measures of conceptual knowledge of Algebra (Genareo et al., 2016). We will revise assessment items in an effort to improve the reliability of the sub-scores. Developing a measure that yields reliable sub-scores will allow us to evaluate potential bi-directional relations between procedural and conceptual knowledge (Rittle-Johnson, Schneider, & Star, 2015).

References

Adelman, C. (2006). The Toolbox Revisited: Paths to Degree Completion From High School Through College (pp. 223). Washington, D.C.: Office of Vocational and Adult Education, Washington, DC.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). Standards for educational and psychological testing. Washington, DC: American Educational Research Association.

Blöte, A. W., Van der Burg, E., & Klein, A. S. (2001). Students' flexibility in solving two-digit addition and subtraction problems: Instruction effects. Journal of Educational Psychology, 93, 627-638. doi: 10.1037//0022-0663.93.3.627

Dowker, A. (1992). Computational estimation strategies of professional mathematicians. Journal for Research in Mathematics Education, 23, 45-55. doi: 10.2307/749163

Genareo, V. R., Deleeuw, W. W., Foegen, A. M., Dougherty, B. J., & Olsen, J. (2016, April 12). Students' understanding of functions and algebra. Paper presented at the 2016 annual meeting of the American Educational Research Association. Retrieved July 29, 2016 from the AERA Online Paper Repository.

Hill, H.C. & Shih, J. (2009). Research commentary: Examining the quality of statistical mathematics education research. Journal for Research in Mathematics Education, 40, 241 – 250.

Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K. C., Human, P., Murray, H., . . . Wearne, D. (1996). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. Educational Researcher, 25, 12-21. doi: 10.2307/1176776

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.

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, 297-312.

National Mathematics Advisory Panel. (2008). Foundations of success: The final report of the National Mathematics Advisory Panel. Washington, DC: U.S. Department of Education.

Rittle-Johnson, B., Schneider, M., & Star, J. R. (2015). Not a One-Way Street: Bidirectional Relations Between Procedural and Conceptual Knowledge of Mathematics. Educational Psychology Review. doi: 10.1007/s10648-015-9302-x

Schmidt, W. H., McKnight, C. C., Cogan, L. S., Jakwerth, P. M., & Houang, R. T. (1999). Facing the consequences: Using TIMSS for a closer look at US mathematics and science education. Dordrecht: Kluwer.

Schneider, M., Rittle-Johnson, B., & Star, J. R. (2011). Relations among conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. Developmental Psychology, 47, 1525-1538. doi: 10.1037/a0024997

Star, J. R., & Newton, K. J. (2009). The nature and development of expert's strategy flexibility for solving equations. ZDM Mathematics Education, 41, 557-567. doi: 10.1007/s11858-009-0185-5

Star, J. R., Newton, K., Pollack, C., Kokka, K., Rittle-Johnson, B., & Durkin, K. (2015). Student, teacher, and instructional characteristics related to students' gains in flexibility. Contemporary Educational Psychology, 41, 198-208. doi: 10.1016/j.cedpsych.2015.03.001

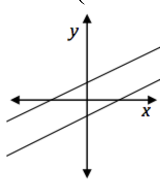
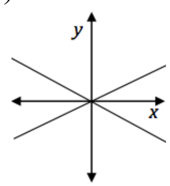
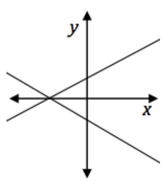
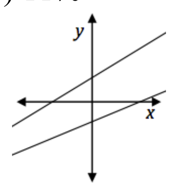
Thorndike, R. M., & Thorndike-Christ, T. (2010). Measurement and evaluation in psychology and education (8th ed.). NY: Pearson.

Table 1. Reliability and Validity Information for Overall Assessment and Item Type Sub-scores

	Cronbach's Alpha	Correlation with Course Grades	Correlation with state test scores
Overall	.85	.51	.57
Conceptual	.66	.42	.47
Procedural	.79	.47	.40
Procedural Flexibility	.66	.34	.48

Note. All correlations are significant at the .001 level except for the correlation between procedural flexibility and course grades, $p < .01$.

Figure 1. Example Items and Performance Data by Item Type

Item Type	Example Easy Item & Response Rates	Example Hard Items & Response Rates
Conceptual Knowledge	Which of the following graphs could represent a system of equations with no solution? A) 79% (correct)  B) 4%  C) 7%  D) 11%  	Look at this pair of equations. **Without solving the equations,** decide if these equations are equivalent (have the same answer). $34=8(x+1)+6(x+1)$ $34=14(x+1)$ A) YES (same answer) 47% (correct) B) NO (different answer) 38% C) CAN'T TELL without doing the math 13%
Procedural Knowledge	Kelley drew a line passing through the point (-1, 4) and (1, 2). What is the y-intercept of the line she drew? A) b = -1 13% B) (0, 3) 68% (correct) C) (3, 0) 13% D) (0, 5) 1%	Which of the following is a factor of $x^2 + 3xy + 10y^2$? A) $x + 5y$ 37% (correct) B) $x + 2y$ 21% C) $x – 5y$ 21% D) $x – 10y$ 12%
Procedural Flexibility	On a timed test, which would be the BEST way to start solving the equation $(7x + 5)^2 = 64$? A) $(7x + 5)(7x + 5) = 64$ 24% B) $(7x)^2 + 2(7x)(5) + 5^2 = 64$ 13% C) $\sqrt{7x+5}=\sqrt{64}$ 57% (correct)	Which way(s) would be mathematically okay way(s) to start solving the problem $x + 4 = 12$? A) $x + 4 – 12 = 12 – 12$ 8% B) $x + 4 – 4 = 12 – 4$ 32% C) $x + 4 – x = 12 – x$ 13% D) A and B 21% E) A, B, and C 21% (correct)