Development of the Elementary Algebra Concept Inventory for the College Context

Claire Wladis
Borough of Manhattan Community College/Graduate Center at the City University of New York
Kathleen Offenholley, Susan Licwinko, Dale Dawes, Jae Ki Lee
Borough of Manhattan Community College at the City University of New York

This study describes the creation and validation of the first concept inventory for elementary algebra at the tertiary level. A 22-item multiple choice/multiple answer instrument was created through a combination of literature review, syllabus review, and collaboration with instructors. The instrument was then revised and tested for content, construct and concurrent validity as well as composite reliability, using a circular process that combined feedback from experts (mathematicians, instructors, and mathematics education researchers), cognitive interviews with students, and field tests using both classical test theory and item response theory. Results suggest that the inventory is a valid and reliable instrument for assessing student conceptual understanding in elementary algebra, as conceptualized in this study.

Keywords: elementary algebra, conceptual understanding, concept inventory

Elementary algebra and other developmental courses have consistently been identified as barriers to student degree progress and completion. Only as few as one fifth of students who are placed into developmental mathematics ever successfully complete a credit-bearing math course in college (see e.g. Bailey, Jeong, & Cho, 2010). At the same time, elementary algebra has higher enrollments than any other mathematics course at US community colleges (Blair, Kirkman, & Maxwell, 2010). Moreover, groups traditionally underrepresented in higher education and in STEM fields are significantly more likely to be placed into elementary algebra. For example, the National Center for Education Statistics reported that from 2003 to 2009, 51.6% of African Americans and 49.5% of Hispanics were enrolled in developmental mathematics courses in college, compared to only 39.4% of whites (2012).

There is evidence that students struggle in these courses because they do not understand fundamental algebraic concepts (see e.g. Givvin, Stigler, & Thompson, 2011; Stigler, Givvin, & Thompson, 2010). Conceptual understanding has been identified as one of the critical components of mathematical proficiency (see e.g. (National Council of Teachers of Mathematics (NCTM), 2000; National Research Council, 2001), and many research studies have documented the negative consequences of learning algebraic procedures without any connection to the underlying concepts (see e.g. J. C. Hiebert & Grouws, 2007). However, developmental mathematics classes at community colleges currently focus heavily on recall and procedural skills without integrating reasoning and sense-making (Goldrick-Rab, 2007; Hammerman & Goldberg, 2003), often because there is pressure for students to pass standardized exit exams that can be exclusively procedural in nature. This focus on procedural skills, divorced from conceptual reasoning, can create a vicious cycle in which the developmental students most in need of explicit instruction in conceptual understanding do not receive it.

At the same time, no validated assessments currently exist to assess conceptual understanding for elementary algebra in the postsecondary context. As a result, instructors cannot systematically detect which incorrect or underdeveloped algebraic conceptions are impeding student progress, and thus they cannot target instruction to address these conceptions explicitly. For these reasons, our team developed an elementary algebra concept inventory (EACI), which we tested for validity and reliability. Research questions included:

1. To what extent does the concept inventory have content, construct, and face validity?
2. How strong is the composite reliability of the instrument?
3. Does the instrument show concurrent validity in being able to distinguish between students with low versus high levels of conceptual understanding in elementary algebra?

Conceptual understanding

The definition of conceptual understanding (and its relationship with other dimensions of mathematical knowledge, particularly procedural fluency) has been much debated and discussed (e.g. Baroody, Feil, & Johnson, 2007; Star, 2005), with as yet no clear consensus. Conceptual understanding and procedural fluency (as well as other mathematical skills) are strongly interrelated (e.g. J. Hiebert & Lefevre, 1986; National Research Council, 2001). However, it can be important to focus on conceptual understanding explicitly, since without explicit instruction in concepts, students may interpret mathematics as a sequence of algorithms, arbitrarily applied, without understanding (e.g. J. C. Hiebert & Grouws, 2007) and may be unable to correctly apply procedures (e.g. Givvin et al., 2011; Stigler et al., 2010).

This study recognizes the interrelatedness of conceptual understanding with other mathematical skills, and defines it this way: An item tests ***conceptual understanding*** if *logical reasoning* grounded in mathematical definitions is necessary to answer correctly, and it is *not* possible to arrive at a correct response *solely* by carrying out a procedure or restating memorized facts. We define a ***procedure*** as a sequence of algebraic actions and/or criteria for implementing those actions that could be memorized and correctly applied with or without deeper understanding of the mathematical justification. For example, consider the following questions:

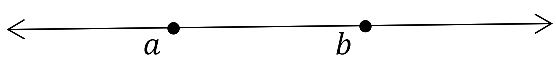
Sample procedural question

If $a < b$, which of the following expressions must ***also*** be true? There may be **more than one** correct answer—select **ALL** that are true.

a. $-a < -b$
b. $2a < 2b$
c. $\frac{a}{2} < \frac{b}{2}$
d. $a - 1 < b - 1$
e. $a + 1 < b + 1$

Sample conceptual question (similar to one from the inventory)

Consider the numbers a and b on the number line below. Which of the following must be true? There may be **more than one** correct answer—select **ALL** that are true.



a. $-a$ is to the left of $-b$ on the number line
b. $2a$ is to the left of $2b$ on the number line
c. $\frac{a}{2}$ is to the left of $\frac{b}{2}$ on the number line
d. $a - 1$ is to the left of $b - 1$ on the number line
e. $a + 2$ is to the left of $b + 2$ on the number line

In the procedural question, it is possible to answer correctly with no understanding of the mathematical reasoning behind the properties of inequality; technically a student could correctly answer this question by using a standard memorized algorithm for manipulating inequalities, even if they do not understand the concepts behind it. For the conceptual question, it is not possible to answer it completely correctly using only memorized procedures without understanding. For example, there is no obvious expression or equation on which to apply procedures. Students could translate the information given by the number line into the inequality $a < b$ and then apply procedures, but this would require conceptual understanding of how to translate between graphical representations of the number line and inequalities.

Conceptualizations of algebra domains in the research literature

There is no one clear consensus about what the core concepts of algebra are. Attempts have been made to categorize algebra by the types of structures or types of actions that are involved (see e.g. Aké, Godino, Gonzalo, & Wilhelmi, 2013; Bell, 1996; Gascón, 1994-1995; Godino et al., 2015; Kaput, 1995; Kieran, 1996; Lee, 1997; Lins, 2001; Mason, Graham, & Johnston-Wilder, 2005; Pinkernell, Düsi, & Vogel, 2017; Rojano, 2004; Smith, 2003; Star, 2005; Sutherland, 2004; Usiskin, 1988), and national standards for algebra have been developed (Common Core State Standards Initiative, 2017; Mathematical Association of America, 2011; National Council of Teachers of Mathematics (NCTM), 2000), although those elementary algebra standards are aimed at K-12 rather than adult learners, which may be problematic since adult learners have been shown in some cases to use different mathematical reasoning from K-12 students (Masingila, Davidenko, & Prus-Wisniowska, 1996; Scribner, 1984).

In order to develop a list of concepts fundamental to elementary algebra at the tertiary level, we conducted a literature review to compile a list concept domains that are both common in research in algebraic thinking and relevant to elementary algebra curricula in the college context (Wladis, Offenholley, Licwino, Dawes, & Lee, n.d.). We began by consulting reviews on the topic (e.g. (Kieran, 2006; Kieran, 2007; Wagner & Kieran, 1989) as well as papers that were cited by, or that cited these reviews. Then, since the last of these larger systematic reviews was published in 2007 (Kieran, 2007), we searched nine mathematics education research journals, nine general education research journals, and five sets of mathematics education conference proceedings, to find any research focused on algebraic thinking; references listed in these papers were also explored. This systematic review (limiting to topics relevant to elementary algebra in the college context) led to a classification of the existing research into four common domains:

- (C1) **Variables and symbolic representation** (see e.g. Bardini, Radford, & Sabena, 2005; Bloedy-Vinner, 2001; Dubinsky, 1991; Furinghetti & Paola, 1994; Knuth, Alibali, McNeil, Weinberg, & Stephens, 2005; Kuchemann, 1978; Malisani & Spagnolo, 2009; Philipp, 1992; Sfard, 1991; Stacey & MacGregor, 1999; Ursini & Trigueros, 2004; Usiskin, 1988)
- (C2) **Equality/Equivalence** (see e.g. Alibali, Knuth, Hattikudur, McNeil, & Stephens, 2007; Falkner, Levi, & Carpenter, 1999; Kieran, 1981; Kieran, Boileau, Tanguay, & Drijvers, 2013; Knuth et al., 2005; Knuth, Stephens, McNeil, & Alibali, 2006; Mevarech & Yitschak, 1983; Rittle-Johnson & Alibali, 1999; Steinberg, Sleeman, & Ktorza, 1990; Zwetzschler & Prediger, 2013)
- (C3) **Algebraic structure sense** (see e.g. Christou & Vosniadou, 2008; Christou & Vosniadou, 2012; Hoch & Dreyfus, 2004; Hoch & Dreyfus, 2005; Hoch & Dreyfus, 2010; Hoch, 2003; Hoch & Dreyfus, 2006; Hoch, 2007; Linchevski & Livneh, 1999; Menghini, 1994; Musgrave, Hatfield, & Thompson, 2015; Novotná & Hoch, 2008; Tall & Thomas, 1991; Thompson & Thompson, 1987)
- (C4) **Functions, proportional reasoning, and covariation** (see e.g. Blanton & Kaput,

2005; Breit-Goodwin, 2015; Carlson, Oehrtman, & Thompson, 2005; Carlson, Jacobs, Coe, Larsen, & Hsu, 2002; Carlson, Oehrtman, & Engelke, 2010; Goldenberg, Lewis, & O'Keefe, 1992; Heid, 1996; Heid & Blume, 2008; Oehrtman, Carlson, & Thompson, 2008; Sitomer, 2014; Thompson, 1994)

In addition, elementary algebra syllabi were also consulted (30 community colleges randomly selected from IPEDS) and participatory action research was conducted with five experienced elementary algebra instructors to outline domains of elementary algebra in the college context and to generate appropriate questions for the inventory (Wladis, Offenholley, Lee, Dawes, & Licwinko, 2017; Wladis, Offenholley, Licwinko, Dawes, & Lee, 2017).

Results and Discussion

Initial Pilot Testing

For an initial pilot test of the instrument, the inventory (V1) was given to a sample of 23 college algebra students at the Borough of Manhattan Community College, City University of New York (BMCC/CUNY) in order to identify problematic questions, issues with wording, and to assess how long it would take students to complete particular problems. Based on these results, the wording of questions was simplified, and some questions were broken into simpler parts. The revised inventory (V2) was given to a group of 160 students who had recently taken elementary algebra; they were asked not only to complete the inventory, but to give feedback on the difficulty and clarity with which problems were posed, as well as to give written explanations of their experiences with the individual questions, and any other feedback that they wanted to share. The time that students spent on each problem was also tracked. Problems that were identified as particularly difficult or unclear or were were missed by a significant number of students, or on which students spent a particularly long time were revisited and revised jointly by the whole group of instructors. Roughly half of the inventory questions was revised at this stage.

Content and Face Validity

In order to assess content and face validity, the revised inventory (V3) was given to 52 instructors who had recently taught elementary algebra at the community college or university level. Instructors came from four different states, many different ethnic/racial and immigrant backgrounds, included faculty who taught at both two- and four-year colleges, and included faculty with a variety of different degree backgrounds. In addition to taking the test, instructors were also asked to give feedback on each question, assessing the clarity and difficulty and giving suggestions for corrections or improvements. They were also asked to assess the topics and concepts covered by the questions and were asked to suggest concepts that were overrepresented, underrepresented, or missing from the current version of the concept inventory. Based on the feedback of faculty, the instrument was revised further.

For example, the following question was included on V3, but was then replaced with other questions in V4 because it was pointed out that it could be answered correctly by applying procedures without understanding:

Which of the following equations are true? There may be **more than one** correct answer—select **ALL** that apply.

a. $2x + 3x = 5x$
b. $2x + 3x = 5x^2$
c. $x^2 + x^3 = x^5$

d. $x^2 \cdot x^3 = x^5$
e. $x^2 \cdot x^3 = x^6$

This question was subsequently replaced with two different questions aimed at testing the conceptual understanding of the algebraic reasoning behind combining like terms and properties of exponents, one of which is similar to the question below:

A student has written: $2x^2(3y-4) + 5x^2(3y-4) = 6x^2(3y-4)$

Which of the following statements is true?

a. The students' work is incorrect because the correct answer is $6x^4(3y-4)$.
b. The student's work is incorrect because the two terms cannot be combined.
c. The student's work is correct because 2 and 5 can be seen as coefficients for the common expression $x^2(3y-4)$.
d. The student can only simplify this expression if they distribute the $2x^2$ and the $5x^2$ first, and then combine like terms.

Reliability, Structure, and Convergent Validity

The resulting instrument was administered to elementary algebra students across four semesters; we report results from the first two semesters (V4 and V5) here. V4 was administered to 28 sections of elementary algebra taught by 23 different instructors, once during the first week of classes (pretest) and once during the last week of classes (posttest). This resulted in 484 completed pretests and 315 completed posttests. Item response theory was used on V4 to assess the extent to which individual items provided good discrimination and were at an appropriate level of difficulty. Items that seemed to be very difficult or very easy, as well as those with negative or low discrimination were revised (V5). V5 was administered to 33 sections taught by 21 instructors, resulting in 431 completed pretests and 192 completed posttests.

Construct Validity

Twenty students who were currently enrolled in elementary algebra were invited to participate in cognitive interviews. Each student was asked to complete the items on their own, and then were asked to participate in a "retrospective think-aloud" interview (Sudman, Bradburn, & Schwarz, 1996), in which they explained what possible approaches to answer the question they considered while completing the questions. Research suggests that concurrent and retrospective think-aloud protocols reveal comparable information, and that retrospective think-alouds may be preferable to concurrent think-aloud protocols because they are less likely to have a negative effect on task performance, especially for complex or challenging tasks with a high cognitive load (see e.g. (Van Den Haak, De Jong, & Jan Schellens, 2003).

Students were then asked to explain what they were thinking about as they considered each question, and then to explain for each answer choice why they did or did not select it. The cognitive interviews focused in particular on several main aims: whether student answers were consistent (i.e. (in)correct thinking yielded (in)correct answers; whether the activities and thinking elicited by the questions was conceptual and in the algebra domain intended by the question; and whether students used test-taking strategies. Cognitive interviews revealed strong consistency of student responses, with almost all students choosing (in)correct answers only when they exhibited evidence of (in)correct reasoning. The vast majority of student responses (including incorrect responses) exhibited some degree of conceptual thinking—for example, even when students were using incorrect reasoning, they were often attending to structural aspects of the expressions or equations on the exam, rather than attempting to apply procedural algorithms to them. There was a low-incidence of use of test-taking strategies.

Reliability and Internal Consistency

Both structural equation modeling (SEM) and item response theory (IRT) were used to assess the reliability of the instrument.

Classical test theory: Using structural equation models for confirmatory factor analysis. Confirmatory factor analysis using structural equation modeling (SEM) was used to model items as predictors of a single latent construct. Average variance extracted (AVE) ranged from 0.76-0.99, indicating very good convergent validity; composite reliability (CR) ranged from 0.99 to 0.9998, indicating excellent reliability (Hair, Anderson, Tatham, & Black, 1998), consistent with requirements for high-stakes testing (Nunnaly, 1978); the standardized root mean square residual (SRMR) ranged from 0.06 to 0.08, suggesting that the model fit was good and supporting the operationalization of the inventory as having a single latent construct (Hu & Bentler, 1999). RMSEA was acceptable, ranging from 0.061-0.064, further supporting the goodness of model fit (MacCallum, Browne, & Sugawara, 1996).

Item response theory. First hybrid three-parameter logit models were explored, with items grouped into three separate groups (based on the underlying probability of randomly guessing that item correctly), each with its own pseudoguessing parameter (Birnbaum, 1968). However, since pseudoguessing parameters were not significantly different from zero, a two-parameter logic model was used. Reliability was assessed using the test information function (tif), where $reliability\ = 1 - \frac{1}{tif(\theta)}$. In IRT, reliability is dependent upon the value of theta, with $\theta = 0$ representing a mean score on the instrument, and other values ta representing the number of standard deviations (SDs) that a score is above or below the mean (e.g. $\theta = -1$ is one SD below the mean). Peak reliability of the test ranged from $\theta = -1.11$ to 0.56, suggesting that the instrument is most reliable for students who are around the mean in algebraic conceptual understanding as measured by this instrument. Both post-tests obtained excellent reliability ($\geq$ 0.9) (Nunnaly, 1978) around the peak, from about 1.5 SDs below the mean to about 0.5 above the mean (see Table 1). Farther away from the peak the reliability remained strong, with acceptable reliability within a minimum of two SDs of the mean on either side on all test administrations (see Table 1). This suggests that the reliability of the test is quite strong for a broader range of students across the ability spectrum. We note that the peak reliability is higher, and overall reliability somewhat stronger, for the posttests than the pretests; this is not surprising, since students have been exposed to algebraic instruction prior to the posttest, but not necessarily prior to the pretest. For example, this may help students to be more familiar with terminology or symbolic representation used on the exam, and that may lead to more reliable results.

Table 1. Reliability values for various values of θ

$\boldsymbol{\theta}$ **values**	**V4 pretest**	**V4 posttest**	**V5 pretest**	**V5 posttest**
Peak reliability	-1.11	-0.52	0.56	-0.44
Acceptable[a] reliability ($\geq$ 0.7)	$[-3.4, 2.1]$	$[-3.7, 2.4]$	$[-2.9, 3.6]$	$[-3.7, 2.5]$
Good[a] reliability ($\geq$ 0.8)	$[-2.4, 0.6]$	$[-2.7, 1.6]$	$[-1.7, 2.6]$	$[-2.9, 1.7]$
Excellent[a] reliability ($\geq$ 0.9)	NA	$[-1.5, 0.5]$	NA	$[-1.4, 0.5]$

[a]Based on the criteria set forth by Nunnaly (1978)

Pre-Test Versus Post-Test Scores

Mean and median scores on the post-tests each semester were not significantly different from the pre-test scores, suggesting that on average, students are likely not gaining any conceptual understanding after one semester of instruction in a traditional algebra class. This is in line with the findings of concept inventories in other subjects, where it was found that some types of instruction could improve aggregate student gain stores, but that on average traditional instruction did not improve outcomes (see e.g. Epstein, 2013; Hake, 1998).

Table 2. Test scores, calculated by percentage of questions answered correctly, for each test administration

	pretest	**95% CI**	**posttest**	**95% CI**
V4	54.2%	[53.3%, 55.1%]	52.9%	[51.6%, 54.2%]
V5	56.4%	[55.5%, 57.3%]	54.2%	[52.3%, 56.1%]

Limitations

We note that this concept inventory is based on a very specific conceptualization of elementary algebra and of conceptual understanding for this subject as well. Because of this, student scores on this inventory reflect only these constructs, and do not necessarily reflect other ways in which conceptual understanding in elementary algebra may be conceptualized.

One major limitation with many previously-developed tests and instruments historically has been that they were developed among a predominately white, middle class population and then were applied to wider populations, including many ethnic minorities and lower-SES students for whom they were not necessarily valid. The City University of New York, which will be used for this research, is highly diverse with an undergraduate population that is roughly 40% first-generation American, more than 75% students of color, roughly 20% first-generation college students, and more than 50% eligible for Pell grants. This diversity makes CUNY an excellent source of information about the validity of the instrument among this population; however, more work with various populations, including rural and suburban populations, is necessary.

Implications

This research demonstrates that it is possible to create valid instruments that can reliably measure some aspects of conceptual understanding in algebra. Some future avenues of research would be to explore further validation of this inventory, e.g., to determine to what extent scores on the EACI differ from scores on validated exams that test procedural fluency in algebra. In addition, future studies that consider which factors correspond to higher versus lower gain scores for whole classes in elementary algebra could help to shed light on which teaching approaches and curricula, etc. can increase student conceptual understanding in elementary algebra.

For practitioners, this study illustrates that it is possible to create questions that target conceptual understanding in elementary algebra specifically. Instructors wishing to create their own questions that would assess this skill could follow some of the process described here and in (Wladis et al., 2017; Wladis et al., 2017). In addition, this study raises important questions about curricula and assessment. Students in this study showed no gains on average in conceptual understanding over the course of one semester of algebra instruction. At the college in which this instrument was tested, all of the learning outcomes on the syllabi are entirely procedural, and as a result, it is likely that most instructors (even those who employ more active learning techniques, of which there were many in this study) do not directly address concepts in their teaching. While this study presents no conclusive evidence of the relationship between these two things, the patterns observed by looking at the pre- and post-test results of the EACI for students in elementary algebra in this study suggest that teaching algebra procedures alone likely does not in and of itself lead on average to gains in conceptual understanding.

References

Aké, L., Godino, J. D., Gonzalo, M., & Wilhelmi, M. R. (2013). Proto-algebraic levels of mathematical thinking
. In A. M. Lindmeier, & A. Heinze (Eds.), *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (2nd ed., pp. 1-8). Kiel, Germany: PME.

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.

Bailey, T., Jeong, D. W., & Cho, S. (2010). Referral, enrollment, and completion in developmental education sequences in community colleges. *Economics of Education Review, 29*(2), 255-270.

Bardini, C., Radford, L., & Sabena, C. (2005). Struggling with variables, parameters, and indeterminate objects or how to go insane in mathematics. *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education,* University of Melbourne, Australia. *2* 129-136.

Baroody, A. J., Feil, Y., & Johnson, A. R. (2007). An alternative reconceptualization of procedural and conceptual knowledge. *Journal for Research in Mathematics Education,* 115-131.

Bell, A. (1996). Problem-solving approaches to algebra: Two aspects. In N. Bednarz, C. Kieran & L. Lee (Eds.), *Approaches to algebra: Perspectives for research and teaching* (pp. 167-185). Dordrecht, The Netherlands: Kluwer.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord, & M. R. Novick (Eds.), *Statistical Theories of Mental Test Scores* (pp. 395-479). Reading, MA: Addison-Wesley.

Blair, R., Kirkman, E. E., & Maxwell, J. W. (2010). Enrollment, Course Offerings, and Instructional Practices in Mathematics Programs at Two-Year Colleges. *Statistical Abstract of Undergraduate Programs in the Mathematical Sciences in the United States Fall 2010 CBMS Survey* (pp. 131)

Blanton, M. L., & Kaput, J. J. (2005). Helping elementary teachers build mathematical generality into curriculum and instruction1. *Zentralblatt Für Didaktik Der Mathematik, 37*(1), 34-42.

Bloedy-Vinner, H. (2001). Beyond Unknowns and Variables - Parameters and Dummy Variables in High School Algebra. In R. Sutherland, T. Rojano, A. Bell & R. Lins (Eds.), *Perspectives on School Algebra. Mathematics Education Library* (22nd ed.,) Springer, Dordrecht.

Breit-Goodwin, M. J. (2015). *Understandings of Proportionality as a Mathematical Structure and Psychological Aspects of Proportional Reasoning in Community College Mathematics Students* (Doctoral dissertation).

Carlson, M., Oehrtman, M., & Thompson, P. W. (2005). Key aspects of knowing and learning the concept of function. *Mathematical Association of America Research Sampler, 9*

Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events: A framework and a study. *Journal for Research in Mathematics Education, 33*(5), 352-378.

Carlson, M., Oehrtman, M., & Engelke, N. (2010). The Precalculus Concept Assessment: A Tool for Assessing Students' Reasoning Abilities and Understandings. *Cognition and Instruciton, 28*(2), 113-145.

Christou, K. P., & Vosniadou, S. (2008). Natural numbers and their interference in students' interpretations of literal symbols in algebra. *International Journal of Psychology, 43*(3-4), 380.

Christou, K. P., & Vosniadou, S. (2012). What Kinds of Numbers Do Students Assign to Literal Symbols? Aspects of the Transition from Arithmetic to Algebra. *Mathematical Thinking and Learning, 14*(1), 1-27.

Common Core State Standards Initiative. (2017). Common Core State Standards Initiative. Retrieved from http://www.corestandards.org/read-the-standards/

Dubinsky, E. (1991). Reflective Abstraction in Advanced Mathematical Thinking. In D. Tall (Ed.), *Advanced Mathematical Thinking* (pp. 95-123) Dordrecht.

Epstein, J. (2013). The Calculus Concept Inventory-Measurement of the effect of teaching methodology in mathematics. *Notices of the American Mathematical Society, 60*(8), 1018-1027.

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

Furinghetti, F., & Paola, D. (1994). Parameters, unknowns and variables: a little difference. *Proceedings of the XVIII International Conference for the Psychology of Mathematics Education, 2* 368-375.

Gascón, J. (1994-1995). Un nouveau modèle de l'algèbre élémentaire comme alternative à l'«arithmétique généralisée». [A new model of elementary algebra as an alternative to "generalized arithmetic"] *Petit X, 37*, 43-63.

Givvin, K. B., Stigler, J. W., & Thompson, B. J. (2011). What community college developmental mathematics students understand about mathematics, Part II: The interviews. *The MathAMATYC Educator, 2*(3), 4-18.

Godino, J. D., Neto, T., Wilhelmi, M., Aké, L., Etchegaray, S., & Lasa, A. (2015). Algebraic reasoning levels in primary and secondary education. *CERME 9-Ninth Congress of the European Society for Research in Mathematics Education* (pp. 426-432)

Goldenberg, P., Lewis, P., & O'Keefe, J. (1992). Dynamic representation and the development of a process understanding of function. In G. Harel, & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy* (MAA notes no. 25 ed., pp. 235-260). Washington, DC: Mathematical Association of America.

Goldrick-Rab, S. (2007). What Higher Education Has to Say about the Transition to College. *Teachers College Record, 109*(10), 2444-2481.

Hair, J. F. J., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate Data Analysis* (5th ed.). Upper Saddle River, New Jersey: Prentice Hall.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*(1), 64-74.

Hammerman, N., & Goldberg, R. (2003). Strategies for Developmental Mathematics at the College Level. *Mathematics & Computer Education, 37*(1), 79-95.

Heid, M. K. (1996). A technology-intensive functional approach to the emergence of algebraic thinking. In N. Bernarz, C. Kieran & L. Lee (Eds.), *Approaches to algebra* (pp. 239-255) Springer, Dordrecht. doi:https://doi.org/10.1007/978-94-009-1732-3_18

Heid, M. K., & Blume, G. W. (2008). Algebra and function development. *Research on Technology and the Teaching and Learning of Mathematics, 1*, 55-108.

Hiebert, J. C., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (1st ed., pp. 371-404). New York, NY: Information Age.

Hiebert, J., & Lefevre, P. (1986). In J. Hiebert (Ed.), *Conceptual and Procedural Knowledge: The Case of Mathematics*. Hillsdale, NJ: Erlbaum.

Hoch, M., & Dreyfus, T. (2004). Structure sense in high school algebra: The effect of brackets. In M. J. Høines, & A. B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (3rd ed., pp. 49-56). Bergen, Norway: PME.

Hoch, M., & Dreyfus, T. (2005). Students' difficulties with applying a familiar formula in an unfamiliar context. In H. L. Chick, & J. L. Vincent (Eds.), *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education* (3rd ed., pp. 145-152). Melbourne, Australia: PME.

Hoch, M., & Dreyfus, T. (2010). Developing Katy's algebraic structure sense. In V. Durand-Guerrier, S. Soury-Lavergne & F. Arzarella (Eds.), *Proceedings of the Sixth Congress of the European Society for Research in Mathematics Education* (pp. 529-538). Lyon, France: CERME.

Hoch, M. (2003). Structure sense. *Paper Presented at the Third Conference for European Research in Mathematics Education,* Bellaria, Italy.

Hoch, M. (2007). *Structure sense in high school algebra.* (Unpublished Doctoral dissertation). Tel Aviv University, Israel.

Hoch, M., & Dreyfus, T. (2006). Structure sense versus manipulation skills: an unexpected result. Paper presented at the *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education, 3* 305-312.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55.

Kaput, J. J. (1995). A Research Base Supporting Long Term Algebra Reform? Paper presented at the *Paper Presented at the Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education,* Columbus, OH.

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

Kieran, C. (1996). The changing face of school algebra. In C. Alsina, J. Alvarez, B. Hodgson, C. Laborde & A. Pérez (Eds.), *Eighth International Congress on Mathematical Education: Selected lectures* (pp. 271-290). Seville, Spain: S.A.E.M. Thales.

Kieran, C. (2006). Research on the learning and teaching of algebra. In A. Gutiérrez, & P. Boero (Eds.), *Handbook of research on the psychology of mathematics education* (pp. 11-50). Rotterdam: Sense.

Kieran, C. (2007). Learning and teaching algebra at the middle school through college levels: Building meaning for symbols and their manipulation. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 707-762). Greenwich, CT: Information Age Publishing.

Kieran, C., Boileau, A., Tanguay, D., & Drijvers, P. (2013). Design researchers' documentational genesis in a study on equivalence of algebraic expressions. *Zdm, 45*(7), 1045-1056.

Knuth, E. J., Alibali, M. W., McNeil, N. M., Weinberg, A., & Stephens, A. C. (2005). Middle school students' understanding of core algebraic concepts: Equivalence & variable . *Zdm, 37*(1), 68-76.

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*,297-312.

Kuchemann, D. (1978). Children's understanding of numerical variables. *Mathematics in School, 7*(4), 23-26.

Lee, L. (1997). *Algebraic understanding: The search for a model in the mathematics education community.* (Unpublished Doctoral dissertation). Université du Québec à Montréal,

Linchevski, L., & Livneh, D. (1999). Structure sense: The relationship between algebraic and numerical contexts. *Educational Studies in Mathematics, 40*(2), 173-196.

Lins, R. C. (2001). The production of meaning for algebra: A perspective based on a theoretical model of semantic fields. In R. Sutherland, T. Rojano, A. Bell & R. Lins (Eds.), *Perspectives of school algebra* (pp. 37-60). Dordrecht, The Netherlands: Kluwer.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power Analysis and Determination of Sample Size for Covariance Structure Modeling. *Psychological Methods, 1*(2), 130-149.

Malisani, E., & Spagnolo, F. (2009). From arithmetical thought to algebraic thought: The role of the “variable”. *Educational Studies in Mathematics, 71*(1), 19-41. doi:https://doi.org/10.1007/s10649-008-9157-x

Masingila, J. O., Davidenko, S., & Prus-Wisniowska, E. (1996). Mathematics Learning and Practice in and out of School: A Framework for Connecting These Experiences. *Educational Studies in Mathematics, 31*(1/2), 175-200.

Mason, J., Graham, A., & Johnston-Wilder, S. (2005). *Developing thinking in algebra*. London: Sage.

Mathematical Association of America. (2011). College Algebra Guidelines. Retrieved from http://www.maa.org/sites/default/files/pdf/CUPM/crafty/CRAFTY-Coll-Alg-Guidelines.pdf

Menghini, M. (1994). Form in algebra: Reflecting with Peacock, on upper secondary school teaching. *For the Learning of Mathematics, 14*(3), 9-14.

Mevarech, Z., & Yitschak, D. (1983). Students' misconceptions of the equivalence relationship. In R. Hershkowitz (Ed.), *Proceedings of the Seventh International Conference for the Psychology of Mathematics Education* (pp. 313-318). Rehovot, Israel: Weizmann Institute of Science.

Musgrave, S., Hatfield, N. J., & Thompson, P. W. (2015). Teachers' meanings for the substitution principle. In T. Fukawa-Connelly, N. Engelke Infante, K. Keene & M. Zandieh (Eds.), *Proceedings of the 18th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 801-808). Pittsburgh, PA: Mathematical Association of America.

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and Standards for School Mathematics*. Reston, VA: NCTM.

National Research Council. (2001). *Adding It Up: Helping Children Learn Mathematics*. Washington, DC: National Academy Press.

Novotná, J., & Hoch, M. (2008). How Structure Sense for Algebraic Expressions or Equations is Related to Structure Sense for Abstract Algebra. *Mathematics Education Research Journal, 20*(2), 93-104.

Nunnaly, J. (1978). *Psychometric theory*. New York: McGraw-Hill.

Oehrtman, M., Carlson, M., & Thompson, P. W. (2008). Foundational reasoning abilities that promote coherence in students' function understanding. *Making the Connection: Research and Teaching in Undergraduate Mathematics Education, 27*, 42.

Philipp, R. A. (1992). The Many Uses of Algebraic Variables. *The Mathematics Teacher, 85*(7), 557-561.

Pinkernell, G., Düsi, C., & Vogel, M. (2017). Aspects of proficiency in elementary algebra. In T. Dooley, V. Durand-Guerrier & G. Guedet (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education* (pp. 1-5). Dublin, Ireland: Institute of Education Dublin City University and ERME.

Radford, A. W., & Horn, L. (2012). *Web Tables: An Overview of Classes Taken and Credits Earned by Beginning Postsecondary Students.* (No. ED-07-CO-0104). Washington, D.C.: National Center for Education Statistics.

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.

Rojano, T. (2004). Local theoretical models in algebra learning: A meeting point in mathematics education. In D. E. McDougall, & J. A. Ross (Eds.), *Proceedings of the 26th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (1st ed., pp. 37-53). Toronto, ON:

Scribner, S. (1984). Studying working intelligence. In B. Rogoff, & J. Lave (Eds.), *Everyday Cognition: Its Development in Social Context* (pp. 9-40). Cambridge, MA: Harvard University Press.

Sfard, A. (1991). On the dual nature of mathematical conceptions. *Educational Studies in Mathematics, 22*, 1-36.

Sitomer, A. (2014). *Adult Returning Students and Proportional Reasoning: Rich Experience and Emerging Mathematical Proficiency* (Doctoral dissertation). (Paper 1794).

Smith, E. (2003). Stasis and change: Integrating patterns, functions, and algebra thrughout the K-12 curriculum. In J. Kilpatrick, W. G. Martin & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 136-150). Reston, VA: National Council of Teachers of Mathematics.

Stacey, K., & MacGregor, M. (1999). Learning the algebraic method of solving problems. *The Journal of Mathematical Behavior, 18*(2), 149-167.

Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education, 36*, 404-411.

Steinberg, R. M., Sleeman, D. H., & Ktorza, D. (1990). Algebra Students' Knowledge of Equivalence of Equations. *Journal for Research in Mathematics Education, 22*(2), 112-121.

Stigler, J. W., Givvin, K. B., & Thompson, B. J. (2010). What community college developmental mathematics students understand about mathematics. *MathAMATYC Educator, 1*(3), 4-16.

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco, California: Jossey-Bass.

Sutherland, R. (2004). A toolkit for analysing approaches to algebra. In K. Stacey, H. Chick & M. Kendal (Eds.), *The future of the teaching and learning of algebra: The 12th ICMI Study* (pp. 73-96). Dordrecht, The Netherlands: Kluwer.

Tall, D., & Thomas, M. (1991). Encouraging versatile thinking in algebra using the computer. *Educational Studies in Mathematics, 22*(2), 125-147.

Thompson, P. W. (1994). Students, functions, and the undergraduate curriculum. *Research in Collegiate Mathematics Education, 1*, 21-44.

Thompson, P. W., & Thompson, A. G. (1987). Computer presentations of structure in algebra. In J. Bergeron, N. Herscovics & C. Kieran (Eds.), *Proceedings of the Eleventh Annual Meeting of the International Group for the Psychology of Mathematics Education* (1st ed., pp. 248-254)

Ursini, S., & Trigueros, M. (2004). How Do High School Students Interpret Parameters in Algebra? *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education,4* 361-368.

Usiskin, Z. (1988). Conceptions of school algebra and uses of variables. *The Ideas of Algebra, K-12*, *8*, 8-19.

Van Den Haak, M., De Jong, M., & Jan Schellens, P. (2003). Retrospective vs. concurrent think-aloud protocols: testing the usability of an online library catalogue. *Behavior & Information Technology, 22*(5), 339-351.

Wagner, S., & Kieran, C. (Eds.). (1989). *Research issues in the learning and teaching of algebra*. Reston, VA: National Council of Teachers of Mathematics.

Wladis, C., Offehnholley, K., Dawes, D., Lee, J. K., & Licwinko, S. (n.d.). Development and Validation of the Elementary Algebra Concept Inventory.

Wladis, C., Offenholley, K., Lee, J., Dawes, D., & Licwinko, S. (2017). An instructor-generated concept framework for elementary algebra in the tertiary context. In T. Dooley, V. Durand-Guerrier & G. Guedet (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education*. Dublin, Ireland: Institute of Education Dublin City University and ERME.

Wladis, C., Offenholley, K., Licwinko, S., Dawes, D., & Lee, J. K. (2017). Theoretical Framework of Algebraic Concepts for Elementary Algebra. In T. Fukawa-Connelly, N. Engelke Infante, M. Wawro & S. Brown (Eds.), *Proceedings of the 20th Annual Conference on Research in Undergraduate Mathematics Education*. San Diego, CA:

Wladis, C., Offenholley, K., Licwino, S., Dawes, D., & Lee, J. K. (n.d.). Learning Progressions for Conceptual Understanding in Elementary Algebra, with a Focus on the Tertiary Context: Synthesizing the Research Literature.

Zwetzschler, L., & Prediger, S. (2013). Conceptual challenges for understanding the equivalence of expressions – A case study. In B. Ubuz, C. Haser & M. A. Mariotti (Eds.), *Proceedings of the 8th Congress of the European Society for Research in Mathematics Education (CERME 8 in Antalya 2013)* (Ankara, METU University ed., pp. 558-567)