Getting *From Here to There!* : Testing the Effectiveness of an Interactive Mathematics Intervention Embedding Perceptual Learning

Erin R. Ottmar (eottmar@indiana.edu)
Indiana University and University of Richmond, 28 Westhampton Way
University of Richmond, VA 23235 USA

David Landy (dlandy@indiana.edu), Robert Goldstone (rgoldsto@indiana.edu**),**
and Erik Weitnauer (eweitnauer@gmail.com**)**
Indiana University, 1101 E. 10th Street
Bloomington, IN 47408 USA

Abstract

We describe an interactive mathematics technology intervention *From Here to There! (FH2T)* that was developed by our research team. This dynamic program allows users to manipulate and transform mathematical expressions. In this paper, we present initial findings from a classroom study that investigates whether using *FH2T* improves learning. We compare learning gains from two different instantiations of *FH2T* (*retrieval practice and fluid visualizations*), as well as a control group, and investigate the role of prior knowledge and content exposure in *FH2T* as possible moderators of learning. Findings, as well as implications for research and practice are discussed.

Keywords: mathematical cognition; concepts and percepts; mathematics education; learning sciences

Introduction

Mastering basic algebraic concepts is extremely challenging, and many students never accomplish it (NCES, 2011). Often, math instruction emphasizes memorization of abstract rules (Koedinger & Alibali, 2008). However, algebraic literacy—the fluent construction, interpretation, and manipulation of algebraic notations—involves not just memorizing, but learning appropriate perceptual processes (Kirshner, 1989; Kellman, Massey, & Son, 2010; Landy & Goldstone, 2007; Goldstone, Landy, & Son, 2010).

Algebra learning involves *seeing* expressions and equations as structured objects, and using these patterns to perform mathematics (Landy & Goldstone, 2007, 2010). Although in some cases the visual and perceptual patterns are fairly easy to see, some object-centered transformations are not immediately obvious in traditional instruction, and must be acquired over practice (Braithewaite, Goldstone, van der Maas, & Landy, under review; Landy, 2010). While this perceptual-motor understanding of algebraic forms is a potentially rich source of student understanding, it also stands as a barrier to learning if visual patterning is not taught in a controlled manner (Marquis, 1988).

Learning technologies offer a promising new approach to teaching math that is not possible with traditional instruction (Clements, 1999; Gee, 2003) and can provide an environment that contributes to improved student performance (Samur & Evans, 2012). The National Mathematics Advisory Panel (2008) highlights algebra as an area of special concern, and notes that while "technology-based drill and practice and tutorials can improve student performance…the available research is insufficient for identifying the factors that influence the effectiveness of instructional software" (p. 23-24). Further, approaches that focus on perceptual-motor training have shown substantial promise (Ottmar, Landy, & Goldstone, 2012; Kellman, Massey, & Son, 2010), but are underexplored relative to other technology-based mathematics interventions. It is anticipated that training students to see correct algebraic structures through dynamic transformations may be a promising approach to teaching algebraic ideas. Rigid motion is a powerful perceptual grouping mechanism (Palmer, 1999), and transformations are naturally memorable and easy to acquire, making these natural tools for helping students grapple with algebra.

In this paper, we describe a learning technology intended to help students acquire appropriate perceptual strategies. We present preliminary findings from two classroom studies using a dynamic computer-based visualization method (*From Here to There!)* designed to enhance middle school students' understanding of algebraic concepts and notations. In our approach, we present symbols as tactile objects whose structure can be appreciated through exploration and manipulation. This approach contrasts with interventions designed to wean students away from perceptual patterns (Kirshner & Awtry, 2004), which can be seen as detrimental to understanding (Noguiera de Lima & Tall, 2008).

From Here to There!

From Here to There! (*FH2T*) is a self-paced interactive application that introduces students to mathematical content through discovery-based puzzles. Rather than simply applying procedures by rewriting different expressions, this technology allows students to physically and dynamically interact with algebraic expression elements, providing a potentially powerful source of perceptual-motor experiences. Below we describe the design theory, features, and goals of the program.

Design Theory and Practice

We approached the construction of *FH2T* from an iterative design stance. We built many different versions of the application instantiating several variations of the basic

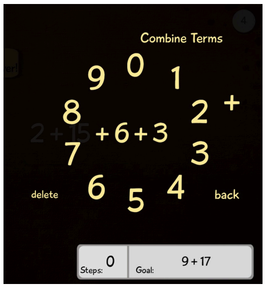

Figure 1: Circular calculator to replace expressions.

equation manipulation interface, the tasks or 'goals' of the user, and the broader application context[1]. For each iteration, we initially evaluated the system with small groups of students. Only the most promising programs were evaluated in classrooms. The experiments reported here reflect a current state, rather than a conclusion, in this process.

Notation manipulation was designed to be as much like a physical environment as possible (Landy, 2010). Transformations are as visually fluid as feasible: elements move smoothly, are picked up by the finger, and dropped. When terms 'split, as in the transformation a*(b+c) => ab+ac, the elements dynamically split.

In order to add or subtract from both sides of an equation, a user taps the equals sign, then is prompted to enter the amount and operation they wish to perform. In certain situations, it is necessary for users to enter numbers that are not in the problem previously (i.e. adding to an equation, or breaking a term into factors). In this case, a calculator with a circular menu is used—numbers, variables, and operations appear in a circle near their targets, and move continuously from the menu location into the equation (Figure 1). In the case of commuting, *a+b* is turned into *b+a* by picking up the *a*, and moving it rightward (or picking up the *b*, and dragging it leftward).

We use a hierarchical structure, with particular *worlds* inside a *universe* (see Figure 2). Each of the 14 worlds cover a particular focal topic, such as 'subtracting multiple terms' and contains a set of about 15-20 *problems*. Locked worlds are presented in black and white on the universe screen, while unlocked worlds are marked in color. Each problem is intended to take between 10 seconds and 1 minute to solve, though the difficulty of particular problems varies considerably. Within each world, problems require users to learn and use new operations alongside previously acquired rules.

Points are used to help a user maintain extrinsic motivation and track their progress (von Ahn, 2013). Participants receive up to 3 points per problem for

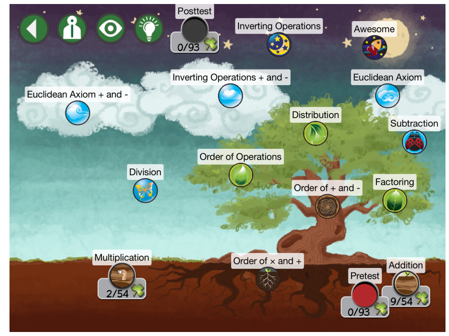

Figure 2: Content Tree Map for *From Here to There!*

,completing it without making calculation or other structural errors, and for completing it in the fewest possible number of steps. Progress is gated by the number of problems completed: at least 14 of 18 problems have to be solved for the user to progress to a new world. This allows a user to avoid extreme frustration by skipping particularly challenging problems, but still requires a fair bit of success at each stage. We also balanced scaffolding user assistance with challenge (Aleven & Koedinger, 2002) by including delayed 'hints' to avoid frustration.

The intelligibility of the goal is also balanced with the richness of flexible and creative mathematical thinking. (Polya, 1954). In many math applications, the user activity seems rote (as in *DragonBox*), or so thoroughly prescribed as to preclude creative thought (as in *Algebra Touch*). *FH2T* uses transformation goals: each problem starts with an equation or expression in a particular form, and states an end state: the user's goal is to transform the equation from the starting form (*here*) to the ending form (*there*). This is intended to help students achieve flexibility in manipulating equations and expressions, compared to having a fixed goal such as "solve for x" (Figure 3). In order to achieve their goal, students perform a series of dynamic interactions, including decomposing numbers (8=5+3 or 11-3), combining terms, applying operations to both sides of an equation, and rearranging terms through commutative, associative, and distributive properties.

The original vision for *FH2T* emphasized visual fluidity. All calculations and transformations were completed automatically: the user initiated the transformation, but the resulting expression simply appeared. However, contrasting approaches suggest that students benefit from being scaffolded through the specific steps required to complete a task in a real-world environment (Tuovinen & Sweller, 1999). Furthermore,

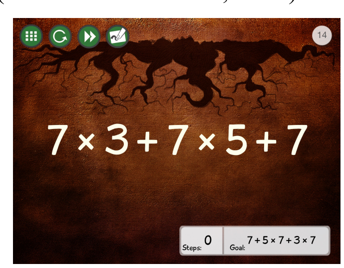

Figure 3: Sample Problem and Goal State in *FH2T*

[1] We began the current effort from an earlier project, *AlgebraTouch (AT)*, which was designed by the second author and Sean Berry in 2007, and has been iteratively improved since then. *AT* has an installation base of approximately 50,000 devices, and a very similar (but not identical) interaction set to that of *FH2T*. *FH2T* branches from a code base of *AT*. We will discuss the equation manipulation interface of *FH2T*, with the understanding that it mirrors in many ways the *AT* system.

teachers frequently expressed concern that students wouldn't learn as well if they did not do the calculations themselves—a perspective compatible with retrieval-based learning (McDaniel & Masson, 1985) and desirable difficulties (Bjork & Bjork, 2011). We take this consideration up explicitly in our study.

Past Research on *From Here to There!*

An initial pilot study was conducted to determine whether *FH2T* contributed to learning gains. 110 6^{th}-8^{th} grade students (41% male, 59% female) from six classes in a large suburban middle school participated in a 4-hour study during six of their regular math periods. All students worked through a series of worlds in *FH2T* at their own pace that covered mathematical topics ranging from addition and multiplication to solving linear equations. In this pilot, a 'reward' system was built into the program. Students first encountered interactions in a *retrieval practice* mode, which required students to recall and enter correct algebraic and arithmetic transformations. After they completed the basic level, they were able to unlock a 'monster level', with especially challenging problems. Once this was completed, students were awarded the *fluid* version of the same interaction, in which correct calculations and transformations were dynamically performed by the system. This was intended to balance the possible benefits of retrieval practice with the minimizing of memory load due to repeated retrievals, and the emphasis on the intended pedagogical domain of algebraic transformations.

Overall, students' mathematical understanding improved 8.5% during the 4 class periods. There was no indication of a floor or ceiling effect: the average accuracy for the posttest problems was 54% (range: 37%-70%). The gains were quite large (effect size=0.40, amounting to one full letter grade) and provide promising results that educational apps, such as *FH2T,* may benefit students when used in combination with classroom instruction. However, due to the non-experimental design of the pilot, we cannot strongly conclude that the learning gains observed were caused from using *FH2T*, per se.

Since students had regular instruction contemporaneously with *FH2T*, it may be that classroom practice led to these gains rather than dynamic interactions. Secondly, *FH2T* ran in two 'modes', which may be differentially responsible for learning gains. When problems were fluidly presented, participants engaged in fast, fluent practice in visual-algebraic patterns. On the other hand, during the initial, retrieval practice phase, participants were forced to engage more explicitly in the specific steps required to solve problems. Observationally, students responded very differently to these two modes. Either of these explanations might plausibly be driving learning gains.

Testing *From Here to There!*

The present study teases these factors apart by dividing participants into three conditions: a business-as-usual control, a *retrieval practice* group, and a *fluid visualizations* group. Using a pre-post design, we aim to differentiate between potential mechanisms behind how *FH2T* produces gains in notation fluency. We also examine the role of content exposure within the *FH2T* program and pretest scores as potential moderators on achievement.

Study Participants and Procedures

Eighty-five sixth and seventh grade students from five classes in a suburban public middle school in the mid-east United States participated in this study during their regular math instruction. All five classes had the same mathematics teacher and students had never had experience using the *FH2T* system.

This study took approximately three hours and occurred over six 30-minute class periods. First, classes were randomly assigned into two groups: intervention (3) and control (2) to ensure that there was not *FH2T* contamination within classes. Next, intervention students were then randomly assigned within classroom to the two intervention conditions (*retrieval practice and fluid visualizations)*. Students in the control classrooms did not use *FH2T*, but received business as usual instruction.

On day 1, all students completed a 30-item pretest that assessed students on procedural facility with various mathematical content. Problems ranged in difficulty from solving basic arithmetic (ex. 3-5+2-3), distributing terms (ex. 3*(5+y+3)), to solving linear equations (ex. 5+y=6+3). Students were presented with a problem and asked to enter their answers using a keyboard. An additional 5 problems asked students to determine whether two expressions were equivalent (ex. does a+b*z+y equal z+y*a+b?).

On days 2-5, students in the intervention classes used *FH2T* to solve problems. The version used in this study was adapted from the pilot in several ways. First, all of the bonus levels were removed. Next, the ordering of the worlds was fixed so that all students had to progress through the content in the same order. Third, the retrieval practice and fluid transformations modes were separated into conditions, so that participants used exclusively one or the other versions of the interaction, embedded within identical problem sets and task space.

Table 1: Pretest and Adjusted Posttest Mean Scores and Standard Deviations as a Function of Condition.

	Pretest		Posttest-Adjusted	
Condition	*M*	*SD*	*M*	*SE*
Fluid Visualizations	9.92	4.1	10.88	0.55
Retrieval Practice	11.00	5.32	10.46	0.62
Control	10.91	3.49	10.46	0.62

In the *retrieval practice* mode, the user first moved the symbols to the appropriate locations to trigger the actions. Next, the user was prompted to write the appropriate resulting subexpression. For instance, if the initial expression was *8-3-2*, and the user tapped the right-hand subtraction, the user would enter "-5", the result of combining *-3* and *-2*. In the *fluid visualizations* mode, the user only had to tap the subtraction sign to initiate the next transformation. The result was a more fluid and dynamic experience since the interface rarely paused.

Students progressed though the worlds at their own pace and completed as much as they could within the time limit. On day 6, all students completed a 30-item post-test assessment, similar in difficulty and form to the pre-test. Pretest and posttests were coded for accuracy and mean scores were calculated for each assessment. It is important to note that the assessment items were designed to measure transfer to problems that are commonly seen in traditional textbooks and worksheets and did not match the transformation goal structure that was presented in the app.

Results

Analysis 1: Do students using *FH2T* improve more than students in a control group?

An analysis of covariance (ANCOVA) was conducted to predict Posttest Scores as a function of Condition, with gender and pre-test scores as covariates. Descriptive statistics for pre-test and adjusted post-test scores are presented in Table 1 and a summary of the ANCOVA results are presented in Table 2. Results show significant differences in post-test scores between conditions, after controlling for gender and pretest $F(2, 84) = 3.61$, $p<0.05$. Next, gain scores were calculated by taking the difference between the pre-test and post-test scores. Significant differences in learning were also found between conditions, $F(2,81)=.4.04$, $p<0.05$ (Figure 3). Post-hoc analyses reveal that students in the fluid visualizations condition (M=2.10) gained more than students in both the retrieval practice (M=-0.22) and control (M=0.22) conditions. No significant differences were found between the control and retrieval practice conditions.

Analysis 2: Does more exposure to content within the *FH2T* app predict improved mathematics performance and learning?

Table 2: ANCOVA of Posttest Scores as a Function of Instructional Condition, With Gender and Pretest Scores as Covariates.

	df	SS	MS	F	η2
Gender	1	31.53	31.53	3.22**	0.04
Pretest	1	1176.47	1176.47	120.01**	0.6
Condition	2	35.37	70.73	3.61**	0.08
Error	80	784.23	90.8		
Total	85	12872			
Corrected	84	2029.65			

One important element to consider is exposure, or how much of the intervention the students actually completed. In this study, we use exposure as a measure of fidelity, to check that greater progress through the program is related to greater performance. We relied on in-app data to create a measure of exposure, calculated as the number of worlds students completed during the duration of the study. On average, students in the fluid condition completed 6 worlds (M=6.04, SD=2.36; addition, multiplication, order of operations + and x, subtraction, division, and order of operations), while students in the retrieval condition only completed the first four worlds (M=4.27, SD=1.97). Students in the control condition, naturally, did not have any exposure to the program (M=0.00).

A hierarchical linear regression was conducted to examine whether increased context exposure within the *FH2T* app predicted posttest performance, above and beyond gender, pre-test, and condition. Dummy codes for the retrieval and control conditions were created to examine whether learning differences remained after adding this additional variable. A significant main effect was found for exposure: for every additional world that the students completed, their posttest accuracy scores increased by 0.76 problems (effect size=0.48) (Table 3). However, after considering exposure in the app, differences between groups were no longer significant.

Table 3: Regression Examining the Contribution of Context Exposure in *FH2T* on posttest scores

	ANCOVA			Main Effect Exposure			Pretest x Exposure		
	B	SE	Beta	B	SE	Beta	B	SE	Beta
Constant	3.83	1.08		0.71	1.38		0.71	1.38	
Pretest	0.88 **	0.08	0.77**	0.72**	0.09	0.63	0.49**	0.13	0.42**
Gender	-1.23 t	0.69	-0.13 t	-1.02	0.65	-0.10	-1.13	0.63	-0.12
Control	-1.77**	0.83	-0.18**	2.97 t	1.62	0.30	2.21	1.60	0.22
Retrieval	-2.20**	0.87	-0.21**	-0.67	0.94	-0.06	-1.20	0.94	-0.11
Exposure				0.76**	0.23	0.48	-0.16	0.43	-0.10
Exposure x Pretest							0.07**	0.03	0.63**
R^2		0.61			0.64			0.66	

Analysis 3: Does prior knowledge moderate the relations between exposure to content and improved mathematics scores?

We examined whether prior knowledge (pre-test scores) moderated the relation between exposure and posttest scores. We tested this by adding the interaction term (exposure x pretest) to the regression. A significant interaction was found (B=0.07, p<0.01) (Figure 4).

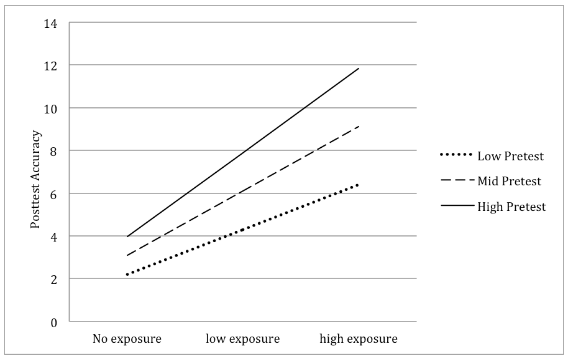

Figure 4: Posttest Scores as a Function of Pre-test Scores and Exposure

Discussion

Overall, we found strong learning gains on the order of one third of a standard deviation from practice with *FH2T*. These gains seem to be primarily due to practice in the more fluid and dynamic version of the application; however, the current approach cannot tease apart whether this is due to a qualitative difference between retrieving explicit rules and perceptual training afforded by the fluid instantiation, or increased topic coverage that the fluid group received. In line with the first possibility, it is notable that participants who practiced retrieval in the application did not show any gains at all, while strong learning gains were found in the fluid visualizations group. These results are fully in line with the theory that algebra literacy comprises strong visual-motor routines (Goldstone, Landy, & Son, 2010; Landy & Goldstone, 2007). With regards to the second possibility, it could be that the additional gains were a result of the fluid condition covering more content that the retrieval practice condition did not get to, but was assessed on the posttest. Future work should manipulate dosage and content exposure that students receive to better understand these effects.

Interventions involving the movement of symbolic forms for algebra learning have been receiving widespread attention in recent years, both in scientific contexts and by the public. Qualitatively, these results—and the strong interest shown by students in solving and discussing problems—suggest promise for tablet-based technologies for teaching abstract algebraic content. This work represents some of the first published outcomes from such perceptual interventions, and may shed light on functional mechanisms. In addition, *FH2T* uniquely focuses on algebraic transformations with a wide variety of initial structures and goal states, attempting to help students think more flexibly about numbers and operations. Transformations of formal algebraic notation is typically demotivating and disengaging for many students; however, students in our studies happily completed several hours of practice, only occasionally becoming bored. One possible explanatory framework for this phenomenon comes from theories of embodiment that suggest that people are intrinsically more engaged when working with their hands (Clark, 2008). Another is that algebra is intrinsically engaging, but that the high cognitive load caused by paper-and-pencil calculations interfere with engaged states. Clearly, much remains to be done.

Although touch-based algebra systems have proved powerful enough to substantially improve algebra skills, notation manipulation is only a small fraction of the important content of algebra. Without connecting to real-world situations, problems, and questions, formal proofs and derivations are largely inert. Further work is currently underway to implement an algebra manipulation system in JavaScript capable of interacting with rich graphics, figures, charts, and text in an html5/canvas webpage.

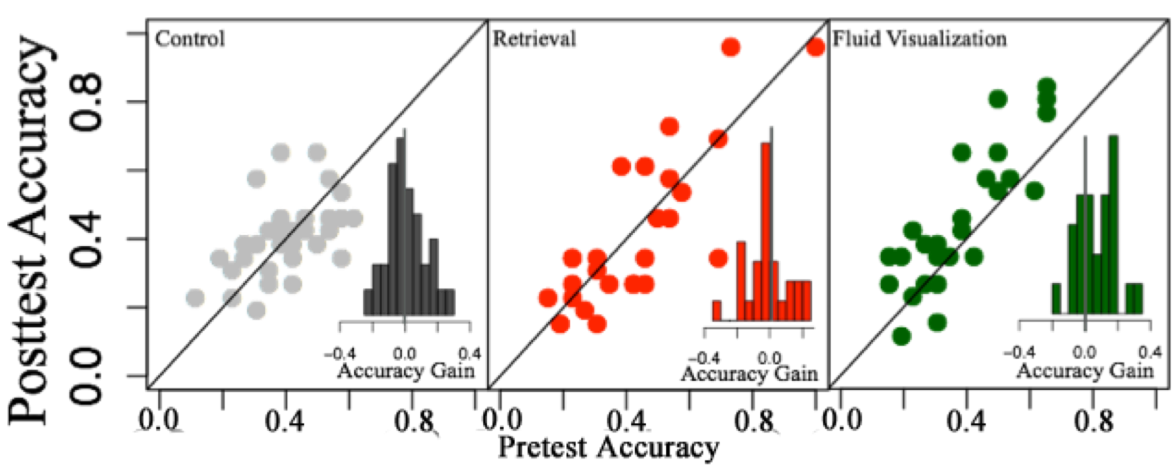

Figure 3: Pre-Test and Posttest Scores and Gains by Condition

Acknowledgments

This work was funded through a grant awarded by the Institute of Education Sciences, US Department of Education (Grant # R305A1100060). We would also like to thank Jessica Lesky, Thad Martin, Catalina Cupanasoiu Jamie Blair, Ryan Smout, Emily Kelly, Megan Delauney, David Brokaw. Aleah Goldin, Taylyn Hulse, and William Barksdale for their contributions.

References

Aleven, V., & Koedinger, K. (2002). An effective metacognitive strategy: learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science, 26*, 147–179.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the real world: Essays illustrating fundamental contributions to society*, 56-64.

Braithwaite, D.W., Goldstone, R.L., van der Maas, D., & Landy, D. (under review). Perception of Structure in Symbolic Arithmetic.

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension: Embodiment, action, and cognitive extension*. Oxford University Press.

Gee, J. P. (2003). What video games have to teach us about learning and literacy. New York: Palgrave/Macmillan.

Goldstone, R. L., Landy, D., & Son, J. Y. (2010). The education of perception, *Topics in Cognitive Science, 2*(2), 265-284. doi: 10.1111/j.1756-8765.2009.01055.

Kellman, P. J., Massey, C. M., & Son, J. Y. (2010). Perceptual learning modules in mathematics: Enhancing students' pattern recognition, structure extraction, and fluency. *Topics in Cognitive Science*, *2*(2), 285-305.

Kirshner, D. & Awtry, T. (2004). Visual salience of algebraic transformations. *Journal for Research in Mathematics Education, 35*(4), 224-257.

Kirshner, D. (1989). The visual syntax of algebra. *Journal for Research in Mathematics Education, 20*(3), 274-287.

Koedinger, K. R., Alibali, M. W., & Nathan, M. M. (2008). Trade-offs between grounded and abstract representations: Evidence from algebra problem solving. *Cognitive Science, 32*(2), 366-397.

Landy, D. & Goldstone, R. L. (2010). Proximity and precedence in arithmetic. *Quarterly Journal of Experimental Psychology*, *63*(10), 1953-1968. doi: 10.1080/17470218.2012.656664

Landy, D. (2010). Toward a physics of equations. In A. K. Goel, M. Jamnik & N. H, Narayanan (Eds.) *Diagrammatic Representation and Inference: 6th International Conference, Diagrams 2010,* Portland, OR, August 2010, (pp. 160-166). New York: Springer.

Landy, D. & Goldstone, R. L. (2007). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(4), 720-733. doi:10.1037/0278-7393.33.4.720

Marquis, J. (1988). Common mistakes in algebra. In A. F. Coxford & A. P. Shulte (Eds.), *The ideas of algebra, K-12: National Council of Teachers of Mathematics Yearbook* (pp. 204- 205). Reston, VA: National Council of Teachers of Mathematics.

McDaniel, M. A., & Masson, M. E. (1985). Altering memory representations through retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(2), 371.

National Mathematics Advisory Panel (2008). *Foundations for Success: The Final Report of the National Mathematics Advisory Panel,* U.S. Department of Education: Washington, DC.

Nogueira de Lima, R. & Tall, D. (2008). Procedural embodiment and magic in linear equations. *Educational Studies in Mathematics*, *67*(1), 3-18.

Ottmar, E., Landy, D. & Goldstone, R. L. (2012). Teaching the perceptual structure of algebraic expressions: Preliminary findings from the Pushing Symbols intervention. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.) *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 2156-2161). Austin, TX: Cognitive Science Society.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology* (Vol. 1). Cambridge, MA: MIT press.

Polya, G. (1954). *Mathematics and plausible reasoning: vol 1: Induction and analogy in mathematics*. Oxford University Press.

Robinson, K. M., & Dubé, A. K. (2009). A microgenetic study of the multiplication and division inversion concept. *Canadian Journal of Experimental Psychology, 63,* 193-200.

Samur, Y. & Evans, M. A. (2012). The effects of serious games on performance and engagement: A review of the literature (2001-2011). Poster session presented at the annual meeting of American Educational Research Association (AERA), Vancouver, British Columbia, Canada.

Tuovinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology, 91*(2), 334-341.

von Ahn, L. (2013, March). Duolingo: learn a language for free while helping to translate the web. In *Proceedings of the 2013 International Conference on Intelligent User Interfaces* (pp. 1-2). ACM.