ABSTRACT

HARRISON, TAYLOR RAY. The Evaluation of iPad Applications for the Learning of Mathematics. (Under the direction of Dr. Hollylynne Lee.)

Algebraic equations have long been a source of confusion and frustration for many students. Many schools are turning towards computers, tablet PCs, and more recently, iPads to assist students' learning in the algebra classroom. The purpose of this research is to review literature on students' learning of algebraic equations, literature on the use of iPads and other tablet PCs in the classroom, literature on digital learning objects, and previously constructed evaluation instruments for digital learning objects. From this review, an evaluation instrument was created to evaluate iPad applications for the learning of mathematics. The instrument created has sixteen criteria in four categories: interactions, quality of content, feedback and support, and usability. This instrument was then applied to seven iPad apps for learning algebraic expressions and equations. The apps were assigned an overall score, and ranked according to this score. A cross-app comparison was conducted to determine which areas might require further developer attention, and where highly-rated apps differentiated themselves from lower-rated apps. Across these seven apps, the scores given in three criteria and one category were found to be significantly correlated with Apple App Store ratings. Using a second rater, the inter-rater reliability of the instrument was tested. Connections to other research involving these apps are made, and limitations of the instrument are discussed.
The Evaluation of iPad Applications for the Learning of Mathematics

by
Taylor Harrison

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Mathematics Education

Raleigh, North Carolina
2013

APPROVED BY:

Dr. Allison McCulloch
Dr. Molly Fenn

Dr. Hollylynne Lee
Chair of Advisory Committee
BIOGRAPHY

Taylor Harrison, raised in Mesic, North Carolina, graduated from the North Carolina School of Science & Math in 2003. A National Merit scholar, he entered the University of North Carolina as a mathematics major before eventually changing his major to computer science and graduating in 2006. After five years in the workforce, math once again called, and in 2011, he entered the masters program in mathematics education at North Carolina State University.

After he graduates in August, 2013, he plans to remain in North Carolina and teach mathematics, hopefully inspiring others as he has been inspired by the many great teachers in his life.
TABLE OF CONTENTS

LIST OF TABLES........................................................................................................... vi 

LIST OF FIGURES........................................................................................................ vii 

CHAPTER 1. INTRODUCTION.......................................................................................... 1 
  The Rise of iPads and Tablet PCs in Schools......................................................... 1 
  Algebra and Equation Solving............................................................................... 2 
  Digital Mathematical Learning Objects................................................................. 3 
  The Need for an Evaluation Metric.......................................................................... 5 

CHAPTER 2. REVIEW OF LITERATURE........................................................................... 9 
  Research on Children's Difficulties in Solving Algebraic Equations.................... 9 
  Research on Digital Learning Objects................................................................... 12 
  Research on Tablets and iPads in Schools............................................................. 18 
  Why Design Matters................................................................................................ 21 
  Previously Constructed Evaluation Models........................................................... 22 
    Multimedia Educational Resource for Learning and Online Teaching (MERLOT)... 23 
    Co-operative Learning Object Exchange (CLOE)................................................ 24 
    Learning Object Review Instrument (LORI)......................................................... 25 
    Le@rning Federation Educational Soundness Specification............................... 27 
    Learning Object Evaluation Instrument (LOEI).................................................... 28 
    Learning Object Evaluation Metric (LOEM)....................................................... 28 
    Identifying Design Principles in Educational Applets (IDEA)......................... 30
Designing Interactive Mathematics................................................................. 32

Digital Tools for Algebra Education.............................................................. 32

Focus in High School Mathematics: Technology Support Reasoning
and Sense Making......................................................................................... 33

Summary........................................................................................................... 34

CHAPTER 3. THE EVALUATION INSTRUMENT...................................................... 36

The Interactions Category................................................................................ 38

The Quality of Content Category................................................................... 43

The Feedback and Support Category............................................................. 45

The Usability Category...................................................................................... 47

Remarks............................................................................................................ 50

Choosing Apps to Evaluate........................................................................... 51

Applying the Instrument & Data Analysis....................................................... 52

CHAPTER 4. RESULTS OF APP EVALUATIONS...................................................... 55

Hands-On Equations for the iPad................................................................. 55

Algebra Touch.................................................................................................. 64

DragonBox+ Algebra....................................................................................... 72

Numerosity: Play with Math!, Chapter 5....................................................... 80

touchyMath...................................................................................................... 86

Cover Up.......................................................................................................... 93

Mathination - Equation Solver....................................................................... 98

Cross-app Analysis.......................................................................................... 104
Correlation with the App Store Ratings

Reliability of the Instrument

CHAPTER 5. DISCUSSION AND CONCLUSIONS

The Criteria

The Apps

Instrument Reliability

Connections to Other Research Involving these Apps

Limitations of the Instrument and Other Areas of Consideration

REFERENCES

APPENDICES

Appendix A. MERLOT Criteria

Appendix B. Co-operative Learning Object Exchange (CLOE) Criteria

Appendix C. Learning Object Review Instrument (LORI) Criteria

Appendix D. Le@rning Federation Educational Soundness Specification Criteria

Appendix E. Learning Object Evaluation Instrument (LOEI) Criteria

Appendix F. Learning Object Evaluation Metric (LOEM) Criteria

Appendix G. Identifying Design Principles in Educational Applets (IDEA) Criteria

Appendix H. Designing Interactive Mathematics Criteria

Appendix I. Digital Tools for Algebra Education Criteria
Appendix J. Focus on High School Mathematics: Technology to Support Reasoning and Sense Making Criteria

Appendix K. The Newly Created Evaluation Instrument
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table 1.</th>
<th>The Criteria Used in the Evaluation Instrument</th>
<th>38</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2.</td>
<td>The Apps Selected for Evaluation</td>
<td>53</td>
</tr>
<tr>
<td>Table 3.</td>
<td>The Evaluation Scores</td>
<td>56</td>
</tr>
<tr>
<td>Table 4.</td>
<td>Pearson Correlations with App Store Ratings</td>
<td>106</td>
</tr>
<tr>
<td>Table 5.</td>
<td>Intraclass Correlation Coefficients</td>
<td>107</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. The equation has been set up on the scale........................................... 58
Figure 2. The pawn pulses as its value is added to the displayed sum.................. 60
Figure 3. Steps required to set up and simplify two similar expressions.............. 62
Figure 4. A typical lesson.................................................................................... 65
Figure 5. The distributive property in action....................................................... 67
Figure 6. A solution gets lost when dividing by (x + 2)........................................ 69
Figure 7. An early level....................................................................................... 72
Figure 8. A few levels later. Several of the cards have changed......................... 73
Figure 9. A level from late in the game............................................................... 74
Figure 10. After multiplying one term by -f, the app indicates where else
the -f card must go............................................................................................ 77
Figure 11. The user must drag the 5 onto the division operator or the 16
onto the unknown............................................................................................... 81
Figure 12. No valid solution is provided for the division....................................... 84
Figure 13. An equation that no longer makes sense........................................... 84
Figure 14. A randomly generated problem......................................................... 87
Figure 15. An interactive lesson on solving linear equations.............................. 88
Figure 16. The user's history of actions after solving a linear equation............... 92
Figure 17. An equation before and after it has been “covered up”...................... 94
Figure 18. A lesson on solving quadratic equations by factoring...................... 98
Figure 19. The hint is not helpful in choosing a pair of factors.................................101
Figure 20. The history is jumbled and difficult to comprehend.................................102
CHAPTER 1

INTRODUCTION

With the increasing popularity of tablet PCs and iPads, scores of K-12 and higher education institutions are adopting these devices to use in the classroom. However, one of the obstacles of using technologies like these to reshape the mathematics classroom is teachers being ill-equipped to determine how best to use these devices (Wilson, 2008). With the hundreds of applications available that could potentially be used in the classroom, not all of them adequately enhance student learning and lead to concept development (Bos, 2009). Teachers are unlikely to use an application in their classroom without some assurance of value and quality (Vargo, Nesbit, Belfer, & Archambault, 2003). Therefore, teachers and administrators need a way to find and select those applications which would most likely be of greatest benefit to the learning of their students. The purpose of this research is to review existing research on the use of mathematical learning objects and tablet PCs in the classroom, review previous criteria for effective learning objects, produce a set of criteria designed to evaluate interactive iPad applications for the learning of mathematics, and apply these criteria to evaluate seven applications for learning algebraic expressions and equations.

The Rise of iPads and Tablet PCs in Schools

Since the introduction of the Windows XP tablet PC in 2002, there has been significant interest in tablet PCs, particularly in education. The engineering schools of both Virginia Tech and the University of Louisville, for example, require all incoming students to purchase tablet PCs (Heib & Ralston, 2010). Murray State University integrates the use of tablet PCs into over 50 different courses. Duke University, Abilene Christian University, the
Rose-Hulman Institute of Technology, Purdue University, and the University of Texas have all established programs distributing iDevices (iPods and/or iPads) to students and faculty members (Hieb & Ralson, 2010; Hlodan, 2010; i.P.O.C., 2006).

The trend is not limited to institutions of higher education. School districts in New York, Illinois, Virginia, California, and North Carolina have collectively spent tens of millions of dollars on iPads alone (Durham Academy, 2012; Hu, 2001; Satariano, 2012). Even kindergarteners are regularly using iPads in the classrooms (Hu, 2011; Subramanian, 2012). And it is not just students in the U.S. that are getting their hands on tablet PCs – in November 2012, the government of Scotland announced a 3-year plan to spend up to £30 million ($47 million) on tablet PCs and netbooks for use in schools, colleges, and universities (Best, 2012).

As of the first quarter of 2013, over 318 million iPhones and 121 million iPads have been sold (APPLinvestors, 2013). By the time a student finishes high school in the U.S., the majority of them will be using technology daily (Franklin & Peng, 2008). Educational theorists (Bereister, 2002; Dede, 2000; Mayberry et al., 2012) advocate that the classroom experience must mirror the complexities of society, integrating technology from students' personal lives into the classroom.

**Algebra and Equation Solving**

Algebra I is typically the high school course with the highest failure rate, as high as 40 to 50 percent in many school districts (Ham & Walter, 1999; Helfand, 2006; Pappano, 2012; Silver, 1995), and students that fail Algebra I are over four times as likely to drop out of high school than those that pass (Orihuela, 2006).
A skill that any successful Algebra I student must acquire is the ability to solve equations. Being able to solve equations requires thinking about numbers and symbols differently than one is likely to have thought about them in previous classes. Filloy and Rojano (1989) describe “cut-points”, the changes the student has to make from their previous ways of arithmetic thinking to new ways of algebraic thinking. The equals sign, for example, is no longer mainly used as an instruction to write down the answer, but is now used regularly to define a relationship. $3+4$ represents not just an instruction to add 3 and 4, but also the result obtained from performing the addition. A student must understand all parts of an algebraic equation – the numbers, the variables, the operators, the equals sign – in order to be able to solve equations in a way that will enhance their algebraic thinking skills.

**Digital Mathematical Learning Objects**

While the iPad itself is relatively new technology (the first edition was released in 2010), interactive digital tools and applications designed to explore mathematical concepts have been around for decades. In recent years, many of these have taken the form of “applets” that appear on the internet. While there are some significant differences between web-based applets and iPad applications (the primary one being the touch interface, rather than typically a mouse and keyboard), they are, for the most part, very similar. Therefore, previous research on learning objects and criteria devised for evaluating them should provide a foundation on which to build a set of criteria for evaluating iPad applications that can be used in the algebra classroom.

Recent recommendations from the National Council of Teachers of Mathematics ([NCTM], 2000) and the Common Core State Standards (National Governors Association
Center for Best Practices, 2010) both call for the use of technology tools in the classroom. Additionally, there has been a call for a more student-centered, task-based learning environment (NCTM, 2000; National Research Council, 2001). Learning objects on the iPad can help educators achieve both of these goals. The iPad provides a safe student-centered environment, conducive to exploration and active learning, in which all students can participate (Cromack, 2008; Loch, Galligan, Hobohm & McDonald, 2011). In addition, teaching practices conducive to a student-centered, task-based learning environment have been found to be more prevalent in classrooms with tablet PCs, and teachers that use tablet PCs report that it is easier to create such environments (Cambre & Hawkes, 2004).

The field of mathematics contains a continuum of objects, ranging from the concrete (e.g., countable objects, visible patterns) to the abstract (e.g., formulas, variables, sets). Many mathematical objects can vary in how abstract they are perceived, depending on instruction and presentation. A learning object, in presenting a mathematical object on-screen, has the ability to make concrete objects more abstract and abstract objects more concrete. This allows students to alter mathematical objects that they previously viewed as concrete, enabling the student to extract properties of the object and see relationships between parts of the object or between different objects. Conversely, a student can be presented an ordinarily abstract object, and be able to interact with it in tangible ways (Bos, 2009). In being able to directly manipulate such mathematical entities as variables, function rules, and equations, students obtain access to the tools of mathematics. These manipulations allow a student to link the specific and the general, and can lead to more complex understandings of objects on which the tools operate (Center for Implementing Technology in Education [CITEd], 2011;
Clements, 2000; Heid, 2003; Lester, 2000; Moyer, Niezgoda, & Stanley 2005). According to Connell (2001), manipulating objects is the first and foundational step in object reification. In order to truly understand a mathematical concept, its mathematical object must be manipulated and conjectures made, then the object manipulated to test these conjectures, leading to further conjectures and problem-solving.

The Need for an Evaluation Metric

There are thousands of mathematics based apps available for the iPad, either for free or for a small price, with new apps continually being added (Attard & Northcote, 2011). As of April 2013, over 600 iPad apps in the iTunes App Store match the keyword “algebra” (Apple, 2013). It is unlikely that a teacher would use these apps extensively in the classroom without some assurance of quality (Vargo et al., 2003). Thus, an effective evaluation metric could greatly reduce the search time for those teachers that only want to examine highly-rated apps (Koppi, Bogle, & Bogle, 2005). Especially for those apps with a cost to purchase, a lack of an evaluation metric could potentially discourage users, particularly those trying to limit costs, to try out an app that might otherwise be an effective teaching or learning tool (Downes, 2002). Teachers might then be compelled to use an unproven app, or create one of their own, rather than using one that has previously been an effective learning tool. Thus, one of the main benefits of learning objects, reusability, is compromised (Kay & Knaack, 2008a).

Choosing appropriate and pedagogically sound software is of great importance (Wilson, 2008, NCTM, 2000). Deciding which software is best for the classroom, however, requires an understanding of pedagogical principles that relate to technology use, something that many teachers lack (Gano, 2011; Wilson, 2008). Teachers must be aware of design
principles and have a means by which they are able to discriminate between those apps which employ pedagogically sound design principles (Bos, 2009; Dick, 2008). Providing these teachers with an evaluation metric based in pedagogical research allows them to do just this, and assures that their criteria for selecting apps is based on qualities that have been backed by research.

An evaluation metric can also be used prescriptively, by designers of apps, during the process of creating the app. There are many designers that aspire to make apps that are useful in the mathematics classroom, but not all of them will be aware of the research that has been done on effective design. Without this knowledge, design decisions that these designers make (sometimes even subconsciously) can hinder the learning of the students whom they wish to help. Giving these designers access to a research-backed set of criteria will allow them to incorporate elements into their apps that will enhance learning, and enable them to make more effective apps (Wiley 2000; Wiley 2002). This, in turn, can introduce multiple perspectives into the mathematics classroom, and give educators a wider range of effective apps to choose from (McCormick, Scrimshaw, Li, & Clifford, 2004).

There is currently a review system in place for iPad apps found in the iTunes App Store. Any user that owns an app can rate the app from 1-star to 5-stars, and optionally write an open-ended review, limited to approximately 5,500 characters. These ratings are then compiled, and a user browsing the store can see the average rating for an app, the distribution of 1-star to 5-star ratings, and read the reviews written by other users (Apple, 2013). But for educators selecting apps to use in the classroom, several shortcomings are present in the system. Typical reviews are open-ended, not addressing specific design characteristics that
might enhance or reduce learning. Most reviewers review only a small selection of apps, making comparing reviews across different apps problematic. Reviews can vary widely in their content, ranging in length from a few words (or simply a numeric rating) to several paragraphs, and likely addressing different aspects of the app, again making comparison of apps difficult. A typically small sample of qualitative reviews makes it difficult to extrapolate experiences to entire populations. Finally, any claims in the reviews that the objects benefited learning are generally not statistically relevant, reliable, or research-backed (Kay & Knaack, 2007; Kay & Knaack, 2009). A standardized evaluation metric, based in research, would alleviate much of these problems.

A handful of instruments have previously been developed to evaluate learning objects, several of which are discussed in Chapter 2. Many of these are not focused on any specific subject area, however, so they fail to address particular aspects that are unique to mathematics. Few, if any, are designed for learning objects on iPads or tablet PCs. By analyzing a number of these instruments as well as examining research in relevant areas, an instrument can be created that is designed specifically for evaluating mathematical apps on the iPad, consisting of criteria that are not only agreed upon by multiple experts, but are also backed by research.

Chapter 2 will consist of a review of literature in three key areas: children's difficulties in solving algebraic equations, the effectiveness of digital learning objects, and the effects of iPads and tablet PCs in the classroom. Several previously constructed evaluation models for digital learning objects were also reviewed. Based on this review, an evaluation instrument was constructed to evaluate iPad applications for the learning of
mathematics, and is presented in Chapter 3. Using this newly constructed instrument, seven iPad applications designed for the learning of algebraic equations and expressions were evaluated and results reported in Chapter 4. Correlation with App Store ratings and inter-rater reliability of the instrument were also calculated and presented in Chapter 4. Based on the evaluations reported in Chapter 4, common weaknesses and strengths of these apps will be discussed in Chapter 5, including areas that may require special attention from developers of these types of apps. In addition, any difficulties in employing the instrument, possible modifications needed to the instrument, and further research that may be needed to refine the instrument will also be discussed.
CHAPTER 2

REVIEW OF LITERATURE

Research on Children's Difficulties in Solving Algebraic Equations

Manipulating and solving algebraic equations has long been a source of difficulty and confusion for students (Filloy & Rojano, 1989; Knuth, Alibali, McNeil, Weinberg, & Stephens, 2005; Knuth, Stephens, McNeil, & Alibali, 2006; Steinberg, Sleeman, & Ktorza, 1990). Many of these difficulties can be traced either to problems in arithmetic or poor or incomplete understandings of algebraic symbols and structures.

The principal part of any equation – the equals sign – is also perhaps the part most often misunderstood by students. Many students view the equals sign as an instruction to evaluate the expression before it (Falkner, Levi, & Carpenter, 1999), or as an announcement that the answer follows (Kieran, 1981). This “operational” or “arithmetical” view is likely not problematic in elementary mathematics, where problems are typically of the form $3 + 5 = ?$. Even with early algebraic equations such as $2x + 10 = 18$ or $x/5 = 4$, the student can invert or “undo” the indicated operations to arrive at the correct value of the unknown. But solving equations such as $2x + 12 = 5x + 3$ requires procedures from outside the domain of arithmetic. Filloy and Rojano (1989) contend that a student must first assign meaning to such an equation, and this, in turn, requires a change in the concept of the equals sign and of the equality of expressions. This change from an operational view to a relational view of the equals sign does not occur spontaneously, however, upon first encounter with such an equation. Intervention and support at this time of transition may be crucial for the learning of students (Filloy & Rojano, 1989). In a study of 177 sixth-, seventh-, and eighth-graders,
Knuth et al. (2006) found that students who expressed a relational view of the equals sign performed better on solving equations of the form $Ax + B = C$ than those students who expressed an operational view of the equals sign, even after controlling for grade level and mathematical ability. In a similar study, Knuth et al. (2005) found that students who expressed a relational view of the equals sign were better able to recognize that equations that had been transformed by performing identical operations to both sides (a common step when solving algebraic equations) maintained the same solution.

A common source of difficulty for children when solving algebraic equations is problems with notation and convention. For example, when asked to find the value of $4 + 3y$ when $y = 2$, many students will evaluate $4 + 32$ to get an answer of 36 (Booth, 1988; Chalouh & Herscovics, 1988). Christou, Vosniadou, and Vamvakoussi (2007) found that the majority of eighth- and ninth-graders indicated that variables and variable expressions could only stand for certain numbers. For example, when asked which numbers in a list could be assigned to the expression $-b$, 52.9% chose only all the negative numbers, while another 11.8% chose only negative whole numbers. Other common errors stem from a lack of notational precision. When learning division, many students assume (or are taught) that the smaller number always “goes into” the bigger number. And so, when dividing $12$ by $3$, it matters little whether the student writes $12 \div 3$ or $3 \div 12$, as long as the answer is 4 (Booth, 1988). Of course, in algebra, the difference between $a \div b$ and $b \div a$ is crucial. Similarly, many students lack the knowledge that the value of $a + b \times c$ is different if you perform the multiplication first than if you perform the addition first. As a result, these students fail to see a need for parentheses in an expression such as $a \times (b + c)$ (Booth, 1988).
All of these notation and convention errors can be problematic when attempting to solve an algebraic equation.

Children's use of informal methods can cause problems when faced with algebraic equations. In a study involving top-achieving sixteen-year-olds, Ekenstam and Nilsson (1979) found that although 82% of them correctly solved the equation $30/x = 6$, only 48% correctly solved the similar equation $4/x = 3$. In the first example, students were able to use the informal method of inspection, a strategy that is not as easily applied to the second equation. In a more recent study, Linsell (2009) examined the correlation between strategy used and success rate of students solving various types of algebraic equations. He found that students who used strategies such as transformations and working backwards were more successful than students using strategies such as guess-and-check and counting strategies, particularly when faced with more difficult equations.

Although research has shown these difficulties and misconceptions are widespread, there is encouraging evidence that they can be quickly overcome or even avoided completely. Hewitt (2012) examined a group of nine- and ten-year-olds that had previously had no introduction to formal algebraic notation. Through the use of Grid Algebra, a computer software designed to support students in learning early algebra, these students became confident in manipulating complex linear equations and appeared to avoid or quickly overcome many of the common difficulties previously discussed, even with little to no intervention from a teacher. Hewitt hypothesizes that this is due to the formative feedback provided by the software, as well as the fact that the software frequently delays computation
of arithmetic expressions, instead treating them as mathematical entities that can themselves be operated on.

Knowing the source of many children's difficulties in solving equations – namely, poor understanding of the different parts of the equation, and adopting informal methods to solve them – is a key step in creating instructional materials to help overcome these difficulties. What follows is a review on literature on a specific type of instructional material, the digital learning object, and the effects they have on students and their learning.

**Research on Digital Learning Objects**

Research has shown that learning objects can enhance active learning and promote collaboration (McCormick et al., 2004; Parish 2004), both of which have been shown to improve student learning outcomes in many cases (Prince, 2004). Clements (2000) found that when using learning objects, students were more likely to cooperate and participate in positive discussion, increasing their social skills and promoting positive social interactions. Even when students disagreed, they tended to disagree about ideas, and were more likely to come to a successful resolution to their disagreements through discussion with their peers. In addition, students are more engaged when using learning objects (Kay & Knaack, 2007), and mathematical problems that would otherwise lack meaningful context can be transformed into tangible situations where students are eager to use their problem-solving skills to find solutions (McCormick et al., 2004; Reeuwijk, 2004).

Students have exhibited positive attitudes towards learning objects, both in anticipation of using them (CITEd, 2011), and after having used them. Students reported that they are easy to use (CITEd, 2011; Kenny, Andrews, Vignola, Schilz, & Covert, 1999;
Reimer & Moyer, 2005), fun and motivating (CITEd, 2011; Kenny et al., 1999; Reeuwijk, 2004; Reimer & Moyer 2005; Vogel, Greenwood-Ericksen, Cannon-Bowers, & Bowers, 2006), more efficient than pen-and-paper methods (Reimer & Moyer, 2005), made them feel more confident about their math skills (Reeuwijk, 2004), and enhanced their learning (Reimer & Moyer, 2005; Riconscente, 2011). In addition, students reported that they appreciated the immediate feedback provided by the learning objects (Reeuwijk, 2004).

As computers became affordable and ubiquitous, attention from schools shifted from “learning to use computers to do math” to “using computers as an aid in a math lesson.” Early computer applications considered the computer to be another display medium and a source of drill-and-practice exercises. However, educators, especially those holding constructivist views of learning, began to oppose this approach and moved to make the computer a tool to aid in student-centered explorations of concepts and open-ended tasks (Durmuş & Karakirik, 2006). Learning objects can provide these learner-centered environments (Parrish, 2004, Loch et al., 2011). Such an environment allows and encourages the student to develop and explore their own mathematical ideas. A student can take control of his or her own learning, making and easily testing conjectures (Clements, 2000). In addition, the affordances of technology allows quicker testing of these ideas and more immediate feedback and apparent results, enabling the student to revisit their conjectures, analyze and revise them, and test them again (Bos, 2009; Laborde, 2007). Students are then more creative and get more self-esteem (Reeuwijk, 2004). Particularly for problems with multiple strategies and entry points, students can work at their own level of thinking, and thus individual differences in students are better addressed (Bos, 2009; Reeuwijk, 2004).
When a student is feeling confused, learning objects can often provide hints and feedback, something that could otherwise not be done without teacher assistance (CITEd, 2011).

Research has suggested that such student-centered learning environments result in deeper understanding of mathematical concepts (Loch et al., 2011). Studies on learning objects suggest that this is indeed the case when learning objects are used in the mathematics classroom. Using learning objects to investigate problems and design solutions has been shown to improve a student's understanding of both the process and the content (Klopfer, Yoon, & Rivas, 2004). In addition, students develop their algebraic thinking and problem solving skills when using learning objects (Reeuwijk, 2004). For example, while working with third graders using a “Pan Balance” applet, Polly (2011) found that students were able to analyze their solutions, evaluate both solutions and equations, and create new equations, simultaneously promoting deeper mathematical understanding as well as higher-order thinking skills. The visual and interactive features of learning objects allow a student to make connections between the visual and the symbolic, making the mathematics easier to understand and benefiting the student's learning (CITEd, 2011; Clements 2000; Reeuwijk, 2004; Suh & Moyer 2007). Because of this focus on the student, the large range of possible actions, and the helpful feedback that can be provided, Laborde (2007) reasons that learning objects are well-suited to help students develop the knowledge underlying optimal solving strategies.

In addition to promoting better understanding of mathematical concepts, learning objects can help a student communicate their understanding (or lack of understanding) to others by providing a window into a student's mathematical thinking. Difficulties and
misconceptions that might otherwise get hidden in traditional methods are more readily revealed when using learning objects (Clements, 2000). Students have to translate their intuition into language and actions that the learning object understands, which in turn, causes the students to use formal mathematics language much more often when describing their ideas to their teachers and peers (Clements, 2000; Hoyles, Healy, & Sutherland, 2008). In addition, learning objects can be helpful for students with language difficulties, including English-language learners, who often have trouble communicating their ideas in the mathematics classroom. A learning object may help them clarify their thinking and demonstrate it to others. Also, several learning objects have support for multiple languages (CITEd, 2011).

The computing power of the technology behind the learning object also offers benefits for students' learning. Rather than the student being bogged down in arithmetic and procedures tangential to the intended learning goal, the computations can be passed off to the technology. This allows the student to focus on the mathematical concepts and models (Reeuwijk, 2004). In addition, the computing power allows for more immediate and accurate calculations, production of graphs, and other visible feedback in response to the user's actions, which has been shown to be valuable for mathematics learning (Forster 2006; Reeuwijk, 2004). The combination of a student-centered environment and the affordances that computing power provides ultimately leads to more efficient learning and cognitive material intake (Pange, 2003; Patsiomitou, 2008; Vogel et al., 2006).

There are practical advantages to learning objects, as well. Most web-based applets are free and iPad/iPhone applications generally range from free to a few dollars. The cost and
effort to produce and distribute learning objects is lower than the alternatives, as well (Duval, Hodgkins, Rehat, & Robson, 2003), resulting in more perspectives and a larger variety of learning objects (McCormick et al., 2004). The content is more readily available, in that any classroom or home with an internet connection and a tool to run the learning object (e.g., computer, iDevice) can have access to any one of countless learning objects in minutes (Duval et al., 2003). Content, whether it be student work, custom problems and assignments, or learning objects themselves, can often be saved and easily shared between students and teachers, or between one teacher and another, either in nearby classrooms or around the world (Duval et al., 2003). This can lead to increased collaboration among teachers and propagation of new ideas about instruction (Parish, 2004).

Learning objects can have a wide variety of applications in the mathematics classroom and can be used with a wide variety of students. Moyer-Packenham, Salkind, and Bolyard (2008) found that among K-8 teachers that used learning objects (specifically, virtual manipulatives), 45% used them in open-ended investigations and problem solving-activities. 37% used them as tools for skill solidification, 14% for introduction of new concepts, and the remaining 4% for other activities such as games, remediation aids, teacher models, and extension of concepts for students achieving above grade-level. In addition to these uses, Haughey and Muirhead (2005) claim that they can also be used to support new types of learning opportunities not available in a classroom environment, extend learning by providing new means for presenting curricular material, and illustrate concepts that are less easily explained through traditional teaching methods.
Learning objects can be effectively used with students of all grades, ranging from kindergarten (Moyer et al., 2005) to high school seniors (Bos, 2009; Kay, 2007). Additionally, learning objects have been shown to have an equally positive effect on males and females, both in regards to student performance (Kay & Knaack, 2008b) and attitude towards the learning object (Kay, 2007; Kay & Knaack, 2008b). And although self-reported computer comfort was shown to be correlated with reception of the learning objects (Kay, 2007), actual performance gains after using the objects was not correlated with computer comfort (Kay, 2007; Kay & Knaack, 2008b). Finally, learning objects have been shown to benefit students in honors classes (Reeuwijk, 2004), as well as those students classified as “low-achieving” (Bos, 2009).

Perhaps of most importance and interest, though, is that learning objects have been shown, in several studies, to increase student performance. In a study in Singapore (Looi & Kim, 2009), for example, a Secondary One class (approximately equivalent to seventh grade in U.S. schools) of 34 students worked for four weeks with a computer applet designed to assist students in constructing algebraic equations from word problems. Students' learning outcomes were measured with a post-test evaluating students' algebraic methods and workings. Results on this post-test were compared with those from a control classroom, which also had 34 students, closely matched in academic grades and previous scores on standardized tests. The post-test scores of the experimental group were significantly higher (p<.001) than the control group. The magnitude of the difference in scores was determined to be very large (eta-squared = 0.23).
In a study examining eighth and ninth grade algebra I classes across four secondary schools in the Netherlands, van Reeuwijk (2004) found that after integrating web applets into the instruction for a chapter, students scored better on the end-of-chapter tests than students without the applets in previous years. No information on significance or size of the improvement was provided from the study, however. In the United States, 95 low-achieving students in 19 classes across two districts participated in a study (Bos, 2009) examining the effects of computer-assisted instruction on a student's mathematical achievement when studying quadratic equations. The students were divided into a control group (47 groups in 9 classes) and an experimental group (48 students in 10 classes), each receiving 55 minutes of instructions per day for eight days. The experimental group's instruction took place in the computer lab, using the Texas Instrument InterActive environment, with lessons that focused on manipulating on-screen mathematical objects. The control group stayed in a traditional classroom, receiving instruction with more of an emphasis on lecture, taking scripted notes, and drill and practice. Post-test scores of the experimental group was significantly greater (p<.001) with a large effect size of .64.

**Research on Tablets and iPads in Schools**

Like desktop computers, early uses of tablet PCs and iPads in the mathematics classroom were primarily as convenience tools or instructional aids, rather than tools for exploring mathematical concepts (Hieb & Ralston, 2010; Loch et al., 2011; Oliver, 2005). In addition, much of the early research on tablets in the classroom focused on university classrooms (e.g., see Fister & McCarthy, 2008; Oliver, 2005; Stickel, 2008). Over time,
however, iPads and tablet PCs started appearing in K-12 schools, and some of the more recent research has taken place there.

Tablet PCs and iPads have been shown to increase student motivation across virtually all grades. In a study of 107 first- and second-graders with low socioeconomic status, all 57 students who solved mathematical tasks using iPad applications indicated positive motivation to use the applications (Segal, 2011). Crichton, Pegler, and White (2012) found that elementary and junior high students demonstrated “great enthusiasm” and described “great satisfaction” for using iPads, after a school year in which iPads were used extensively in the classroom. In a study examining fifth graders using an interactive iPad app to explore fractions, all 122 students reported that the app was fun and that they wanted to use the app again, and 95% reported that they thought their friends would enjoy it (Riconscente, 2011). Studies at the university level have indicated improved motivation (Galligan et al., 2012), increased student attention (Wise, Toto, & Lim, 2006), increased attendance and retention (Romney, 2010), very high participation rates in voluntary tablet activities (Vu, 2007), and increased participation in class (Reba & Weaver, 2007; Romney, 2010).

The use of Tablet PCs can lead to more efficient and effective learning. The use of regular, real-time formative assessment means that students and teachers can become more easily aware of what the student has mastered or of the student's misconceptions. This in turn, allows the teacher to focus on these misconceptions, rather than concepts already mastered by the student, leading to more efficient teaching and learning (Kowalski, Kowalski, & Hoover, 2007; Kowalski, Kowalski, & Gardner, 2009). Tutty and White (2006) argue that the tablet PC classroom environment is more effective than the traditional format,
partly due to the increased emphasis on students' sense-making processes and on the social aspects of learning. Students feel that they are able to learn mathematical concepts well in classes that use tablet PCs (Fister & McCarthy, 2008; Loch et al., 2011). Additional studies and reports suggest that the use of tablet PCs and tablet PC applications can result in an increased mathematical fluency (Attard & Northcote, 2011), students taking ownership of their learning (Fister & McCarthy, 2008), and an overall positive effect on student learning (Wise et al., 2006). As further evidence, Fister and McCarthy (2008) compared exam scores in an Introductions to Proof course at Murray State University (MSU). The class in which all students used a tablet while in class scored 16% higher than the class without tablets, on an identical exam. In Fall 2006, assessments were given to several tablet classrooms at MSU and to a control of traditional classrooms. The tablet classrooms showed a 10-15% improvement in their scores over their non-tablet counterparts. The use of tablet PCs in these classrooms also generated campus-wide discussion and coordination of effective teaching (Fister & McCarthy, 2008).

One of the primary differences between a desktop PC and a tablet PC, besides portability, is the interface. Tablet PCs generally have a touch interface (either using a finger or a pen/stylus), while the most common interface for a desktop PC is a mouse and keyboard. Segal (2011) studied the differences between the effect these two interfaces had on young students' learning of mathematics. One hundred seven six- and seven-year-olds participated in completing two mathematical tasks, some completing the tasks on a tablet PC using the touch interface, and others performing identical tasks on a desktop with a mouse. In the first task (a counting and addition task), students using the touch interface spent significantly (p
< .001) less time solving the problem (24% less, on average), and used more advanced strategies. In the second task (a number line estimation task), students again spent significantly (p < .001) less time solving the problem (32% less, on average). There was no significant different in accuracy between the students who used the tablet PCs and those who used the desktop PCs.

Tablet PCs have practical advantages in the classroom, as well. The touch (or pen) interface allows symbolic and graphical information to be easily written electronically, and enables the teacher to adjust lectures in real-time, to explore different solution paths or in response to students' reactions or feedback (Loch & Donovan, 2006). One of the requirements for an effective Virtual Learning Environment according to the Shared Content Object Reference Model (SCORM) is a consistent standard for running and launching applications (Advanced Distributive Learning, 2009). The iPad in particular has an operating system which satisfies this – all applications are launched from the same place, and folders can be set up so that applications can be manually grouped according to course, topic, or any other desired grouping. Additional advantages include portability, a functional screen size, an abundance of apps, multiple routes for internet access, and multimedia, the combination of which can unleash numerous possibilities for learning and teaching (Gupta, 2010; Kinash, Brand, & Mathew, 2012; Loch et al., 2011).

**Why Design Matters**

Not all learning objects are effective at enhancing students' learning, and using technology for the sake of novelty and interactivity is not enough to meet the increasing standards of today's educators and policy-makers (Bos, 2009). There is a growing recognition...
that outcomes for learning are dependent on software design. The design of technology tools
can drastically affect how students interact with the tools, and this, in turn, may influence
students' understandings of content as well as their acquisition of problem-solving skills (e.g.,
Hoyles & Noss, 2003; Lee & Hollebrands, 2006). Learning objects, through their design, can
also affect the available strategies a student can use, since they can afford or constrain certain
actions (Underwood et al., 2005).

The design of a learning object can be greatly influenced by the designers'
epipistemological and pedagogical beliefs about mathematics, problem solving, and teaching
(Lee & Hollebrands, 2006). Effective design must be rooted in epistemological frameworks,
and is possible only if the designer has a reflexive awareness of these frameworks (Bannan-
Ritland, Dabbagh, & Murphy, 2000).

**Previously Constructed Evaluation Models**

Several authors, researchers, and organizations have devised models for evaluating
learning objects, though not necessarily for the same reasons. Online learning object
repositories, for example, may have criteria that must be met for inclusion into the repository
(e.g., see Le@rning Federation, 2007). Other evaluation models attempt to provide a
quantitative score to a set of existing objects in order to rank them or to assist educators
wishing to use quality learning objects (e.g., see Bokhove & Drijvers, 2010; Multimedia
Educational Resources for Learning and Online Teaching [MERLOT], 2013). Still others are
meant as prescriptive standards intended for designers of learning objects (e.g., see CLOE,
2004; Le@rning Federation, 2007).
Reviewing these evaluation models will provide a good foundation on which to construct an instrument to evaluate mathematical iPad applications. First, some non-mathematics-specific models will be reviewed, in order to glean good design principles for learning objects in general. Then, models specific to mathematics will be reviewed, in order to glean principles that are either specific to mathematics, or perhaps of more importance in mathematical learning objects. The principles gathered from this review will be combined with findings from other research in this chapter to construct the evaluation instrument for the evaluation of mathematical iPad applications.

**Non-mathematics-specific Evaluation Models**

*Multimedia Educational Resource for Learning and Online Teaching (MERLOT)*

MERLOT maintains a repository of approximately 40,000 learning materials across various disciplines. Each discipline (e.g., Math and Statistics) has an Editorial Board consisting of members with (a) expertise in the scholarship of their field, (b) excellence in teaching, (c) experience in using technology in teaching and learning, and (d) connections to professional organizations in their discipline. The MERLOT Management Team creates a framework of evaluation criteria and provides them to the Editorial Board of each discipline. The Editorial Board then applies the criteria to test cases in their discipline, and uses these to develop evaluation guidelines and criteria that are applied to all materials in the discipline. The Mathematics and Statistic set of criteria is identical to the non-discipline-specific framework. The complete set of criteria can be found in Appendix A.

MERLOT's evaluation criteria are separated into three categories: (1) *Quality of Content*, which addresses both validity and educational significance of the content, (2)
Potential Effectiveness as a Teaching-Learning Tool, which addresses such things as alignment with learning objectives and whether the use of technology enhances learning, and (3) Ease of Use, which includes interface and feedback. Each Editorial Board performs a cursory review of learning materials in their discipline to identify worthy candidates for in-depth reviews. At least two members of the Editorial Board then perform reviews using the evaluation criteria and submit them to an editor. From these, the editor creates a single composite review that is then published and attached to the entry for the learning object in the repository for users of the repository to see (MERLOT, 2013).

Co-operative Learning Object Exchange (CLOE)

The (now-defunct) Co-operative Learning Object Exchange (CLOE) was a grant-funded partnership from 2004-2008 between 29 universities across the world to collaboratively develop, adapt, and evaluate learning objects. A set of criteria was created, both to guide university participants in creating learning objects and as a metric for other participants to review those learning objects. The criteria were modeled after the MERLOT model, but consisted of fewer questions, and required only that reviewers choose one of three answers (“not at all”, “somewhat”, or “definitely”) for each question, rather than providing detailed open responses. Criteria fell into three categories, similar to MERLOT's categories: Quality of Content, Effectiveness as a Teaching/Learning Tool, and Ease of Use/Usability. The complete set of criteria can be found in Appendix B.

The review process for CLOE learning objects differed slightly from MERLOT’s. Learning objects first were subjected to a functionality review by the “gatekeeper” to ensure that links work, plug-ins are available, platform and browser compatibility are identified, etc.
Other reviewers were separated into two categories: instructional design experts and subject matter experts. Learning objects had to receive satisfactory reviews (using the aforementioned criteria) from reviewers in both categories before the learning object was published in the CLOE database for other CLOE participants to use (CLOE, 2004).

**Learning Object Review Instrument (LORI)**

The Learning Object Review Instrument (LORI) was developed by Vargo et al. (2003), and later improved by Nesbit, Belfer, and Leacock (2004; 2009). A primary goal of LORI was to balance assessment validity with efficiency of the evaluation process (Leacock & Nesbit, 2007). LORI is noteworthy in that quantitative studies have been performed examining the inter-rater reliability as well as correlation of ratings to learning outcomes of students. Also noteworthy is that LORI has been revised several times, at least in part in response to some of these studies.

Early versions of LORI consisted of ten items. Items included such things as *Content Quality*, *Feedback and Adaptation*, and *Presentation Design* (see Appendix C for the full list), and were each rated on a numerical scale from 0-4 (1-5 in later versions). Each item was accompanied by a description of qualities an item receiving the highest rating would possess, qualities an item receiving the lowest rating would possess, and an example of a hypothetical item that would be rated a 2 (3 in later versions). Later versions reduced the list to nine items (*Aesthetics* was no longer included), and then again to eight items (*Reusability* was no longer included). Other items were slightly renamed and had their descriptions updated from version to version.
Shortly after developing LORI, Vargo et al. (2003) conducted a study to examine inter-rater reliability, both with and without collaboration between the raters. Twelve educational technology professionals and university faculty participated in the study, though only 10 participants’ data were used, due to missing ratings for two participants. Participants received no special training or information about LORI other than the instrument itself, and in most cases, they did not have specific knowledge or expertise matching the learning objects' subject matter.

Using intraclass correlation (ICC) as an indicator for inter-rater reliability and an arbitrary threshold of 0.80 for sufficient reliability, three of the eight items achieved sufficient reliability before collaboration of the participants, and six of the eight items achieved sufficient reliability after collaboration. Two of the LORI items (standards compliance and accessibility) were not considered due to insufficient variation and violation of the assumption of normality. Both of these two items, however, saw majority agreement among raters, both before and after collaboration.

More recently, Akpınar (2008) conducted a study examining the correlation of LORI ratings with teacher usability ratings, student usability ratings, and learning outcomes of students using the learning objects. Usability ratings were based on questionnaires asking the teacher or student about ease of use, and whether they approved of the screen design, text and picture orientation, the interactive mechanisms, and the tools used in the learning object. Student learning outcomes were measured by pre- and post-test scores.

Teacher usability ratings were found to be significantly correlated with six of the nine criteria examined. However, neither student usability ratings nor student learning outcomes
were significantly correlated with LORI ratings, suggesting that LORI ratings alone are not sufficient in predicting significant learning outcomes of users.

*Learning Federation Educational Soundness Specification*

The *Learning Federation* was a five-year development initiative for building a collection of online content for use in Australian universities. The *Learning Federation* Digital Content Repository is the result of this, and provides access to over 3,000 learning objects (over 900 in mathematics), the majority of which are commissioned by the *Learning Federation* itself. The *Learning Federation Soundness Specification* (*Learning Federation, 2007*) was used both to guide content development for learning objects commissioned by the *Learning Federation*, as well as to filter non-commissioned content that is proposed for inclusion in the repository.

The *Learning Federation* sought to produce and select learning objects that enhance student learning (*Learning Federation, 2003*). In order to do this, they based their specifications on research into how students learn and how teachers facilitate this learning. Specific areas of focus for the research they examined included engagement, constructivist learning, inquiry, successful learning, critical pedagogy, and appropriate use of the medium. Through this process, they produced the Educational Soundness Specification, consisting of 39 measures in four principles: *Learner Focus, Integrity, Useability*, and *Accessibility* (the complete list of 39 items can be found in Appendix D). They also created other specifications, including technical requirements and rights management, but maintain that educational soundness is the main determinant in assessing the quality of a learning object.
Learning Object Evaluation Instrument (LOEI)

Haughey and Muirhead (2008) developed the Learning Object Evaluation Instrument (LOEI) to evaluate learning objects specifically for the K-12 sector. Unlike most other evaluation models, LOEI is not meant to provide a comparison between learning objects, but rather a way for a reviewer to determine the value of a learning object in and of itself. LOEI consists of 14 criteria in five categories: Integrity, Usability, Learning, Design, and Values (the complete list of 14 criteria can be found in Appendix E). The criteria were drawn from four sources: 1) CLOE, 2) the Le@rning Federation Soundness Specification, 3) LORI, and 4) criteria developed with respect to the special concerns of the K-12 environment.

To assess the quality of LOEI, the researchers used the instrument to evaluate 22 learning objects, made minor refinements to the instrument, and then at a later time, evaluated 14 more learning objects. Through these evaluations, the researchers agreed that the instrument was easy to use and found a “high degree of congruence” in their assessments (Haughey & Muirhead, 2008).

Learning Object Evaluation Metric (LOEM)

The Learning Object Evaluation Metric (LOEM) was developed by Kay and Knaack (2008a) after a review of literature on instructional design and of existing evaluation models, including CLOE, LOEI, LORI, MERLOT, and a model previously constructed by the researchers. Through this review, Kay and Knaack identified five main constructs for evaluating learning objects: Interactivity, Design, Engagement, Usability, and Content. They noted that most previous models were largely theoretical, without much, if any, quantitative
research on the models or estimates of their reliability and validity. They expressed a desire to systematically assess a multi-component evaluation model for evaluating learning objects.

The original version of LOEM consisted of 29 items in the five aforementioned constructs (the complete list can be found in Appendix F). In evaluating a learning object, each of the 29 items were given a score from 1 to 3. Twelve items were then excluded from the model due to either low correlation with student evaluations, teacher evaluations, and student performance, or due to insufficient fit into principal component analysis. In addition, the *Content* construct was not found to be a significant factor. Kay and Knaack hypothesize that content may have had a negligible influence because teachers may have filtered for content issues before selecting learning objects to use.

In order to assess the reliability and validity of LOEM, Kay and Knaack (2008b) performed a study consisting of 1113 students and 33 teachers in 21 middle and secondary schools using 44 learning unique learning objects, a much larger and diverse sample than previous studies using other evaluation models. The learning objects were assigned LOEM ratings by a team of four teachers after extensive training on the use of LOEM. Students were asked to evaluate the learning objects after using them, via questionnaires addressing three constructs: *learning*, *quality*, and *engagement*. The teachers were also asked to evaluate the learning objects, using a different set of questions as the students, but addressing the same three constructs. In addition, student performance was measured using pre- and post-tests.

Internal reliability for each of the four LOEM constructs was found to be between 0.70 and 0.80, which the researchers noted as moderate, but acceptable. A principal component analysis was performed to explore whether 17 items in LOEM could be grouped
into the four separate constructs that the model proposed. While there was slight overlap in some of the items – for example, “appropriate language level”, while in the Usability construct, was somewhat correlated with the Engagement factor – the researchers determined that overall, the resulting structure fit the evaluation model well. Correlations between the four constructs were significant, but determined to be small enough to support the assumption that each construct was distinct.

Of the four LOEM constructs, three (design, engagement, and usability) were found to be significantly correlated with all three constructs in the student evaluations (learning, quality, and engagement). For the teacher evaluations, both the interactivity and the usability LOEM constructs were significantly correlated with teachers' assessments of quality. The design construct was significantly correlated with teachers' assessments of quality and learning, but not engagement. The engagement construct was significantly correlated to all three constructs in the teachers' assessments. Finally, student performance was examined. Student performance was significantly and positively correlated with each of the four LOEM constructs. This is noteworthy because it is one of the few studies that successfully links an increase in student performance to the ratings of a learning object evaluation model.

**Mathematics-specific Evaluation Models**

*Identifying Design Principles in Educational Applets (IDEA)*

The IDEA group was formed to cull design principles from a set of applets developed for mathematical education during the Educational Software Components of Tomorrow (ESCOT) project. Underwood et al. (2005) details the process the IDEA group undertook and the resulting set of design principles. The IDEA group consisted of members with various
areas of expertise (e.g., teacher, software developer, educational technologist), some of which were involved with the development of the applets in the ESCOT project. They sought to identify design principles that successfully supported problem solving and learner-centered design issues.

Each member individually reviewed each of the 25 selected applets, noting strong and weak characteristics, and selected their five highest ranking and lowest ranking applets, based on overall quality. Then, through discussion with the whole group, an aggregate ranking was formed. The group then generated design principles based on characteristics of the highest ranked and lowest ranked applet. The remaining applets were assigned to smaller groups, where this process was continued to generate additional design principles. The design principles for the small groups were combined, and an aggregate list was created. Finally, the list was reduced to include only those principles which applied to multiple applets. The group decided not to focus on those principles that pertain to ease-of-use, due to the large body of existing literature on the subject of usability. The resulting 21 principles (see Appendix G) were grouped into three categories: motivation, presentation, and support for problem solving.

After producing their list of design principles, the IDEA group also analyzed previously recorded video of students using the applets, to determine whether adherence (or lack thereof) to the design principles promoted or hindered the learning of the students. While their analysis was limited to videos of only two of the applets being used, the group determined that within each video, evidence supporting the design principles was found.
Designing Interactive Mathematics

Designing Interactive Mathematics was published by June Lester (2000), a long-time mathematics instructor and researcher in the field of design of onscreen mathematical technology. While the paper itself is not based on any particular scientific study, the author draws from a variety of sources, including her own experience, to create a preliminary list of potential design principles for the interactivity of a mathematical learning object.

The majority of the article is focused on two principles. First, “if possible, model the behaviour of onscreen mathematical objects by observing and analyzing their “natural” mathematical properties; otherwise, invent and standardize behaviours” (p. 3). This principle encompasses such things as mathematical correctness and parametric independence. Second, “make interactive behaviours visible, at all stages of the interaction” (p. 6). This includes highlighting possible interactions, providing feedback during the interaction, and showing the resulting state of the object after the interaction. Several additional design principles (see Appendix H) are also provided, along with comments on how they can be applied to mathematical learning objects.

Digital Tools for Algebra Education

In preparation for studies into the effect of digital tools on students’ learning of algebra, Bokhove and Drijvers (2010) were faced with the issue of selecting appropriate tools. They designed an evaluation instrument for such tools, based on research in three key areas: algebra didactics (symbol sense), theories on tool use (instrumental genesis), and assessment (feedback and formative assessment). The result of this research was a set of 27 initial criteria (see Appendix I for the full list) grouped into four categories: algebra criteria,
tool criteria, and assessment criteria, along with a fourth category, general criteria, that were more practical in nature, such as the cost of the tool.

Bokhove and Drijvers then sent the list of criteria to 47 experts in the field of mathematics and algebra education, and asked them to rate each criterion on a scale of 1 (unimportant) to 5 (important), with an option to give no answer. The experts were also asked to comment on the thoroughness of the instrument or to suggest additional criteria. Bokhove and Drijvers received 27 responses with rated criteria, and used these to assign relative weights to each criterion. Every criteria scored an “above neutral” average (the lowest rated criteria received an average rating of 3.41, with 3 being “neutral”), and no additional criteria were suggested by the experts. All four categories received approximately equal weightings, with the general category slightly behind the other three. Few other evaluation instruments assign importance to different criterion, and even less do so quantitatively, so Bokhove and Drijver’s efforts are notable in this regard. The five criteria receiving the highest weights (i.e., the ones judged to be the most important by the experts that responded) were stability and performance of the tool, the tool is easy to use for a student, the tool is able to display formulas correctly, the tool is mathematically sound and faithful to the underlying mathematical properties, and the tool stores the answers given by a student.

Focus in High School Mathematics: Technology to Support Reasoning and Sense Making

NCTM’s Focus in High School Mathematics: Reasoning and Sense Making is a series of books designed to help teachers make reasoning and sense-making foundational to the math that is taught to students in high school. The Technology to Support Reasoning and
Sense Making volume (Dick & Hollebrands, 2011) is focused on how to incorporate technology into the classroom to accomplish this goal. In the introduction, Dick and Hollebrands provide “A Guide for Choosing and Using Interactive Technology Scenarios”, in which they present ten questions that teachers should ask when choosing an interactive technology scenario (see Appendix J for the full list). Most questions (e.g., “Is the action purposeful, thoughtful, and deliberate?”) are similar to criteria found in other instruments. While the guide is not explicitly stated to be research-based, the authors had both previously been involved in much research on the topic of technology and student learning (e.g., see Dick, 2008; Lee & Hollebrands, 2006; Underwood et al., 2005; Zbiek, Heid, Blume, & Dick, 2007), and it is reasonable to assume that this guide is substantiated by that research.

Summary

The evaluation instruments reviewed come from a variety of sources. Some instruments were developed by a small team or an individual, while others were developed through collaboration with educators or experts around the world. Some were created through a systematic process of formation and refinement, while others did not have an established method for the creation of criteria. Some instruments were based on years of research, others were based on years of personal experience and anecdotal evidence. Despite all these differences, however, there are several design principles that appear across numerous instruments (sometimes with slightly different phrasing). These common design principles – particularly the ones substantiated by other research in this chapter – will form a set of criteria that will be the basis of the new evaluation instrument that will be constructed
in Chapter 3. These criteria will then be adjusted, and new criteria added, based on research particular to the use of iPads and other tablet PCs for learning.
CHAPTER 3

THE EVALUATION INSTRUMENT

Using these evaluation models and other research on digital learning objects and tablet PCs in the classroom as a guide, we can cull several commonalities and design principles to create an evaluation instrument for iPad apps used for the learning of mathematics. The purpose of this chapter is to present, describe, and justify the evaluation instrument created and used as part of this research.

From across the ten instruments and sets of criteria presented in Chapter 2, criteria that appeared in multiple instruments were identified. Since criteria in different instruments were rarely worded identically, judgment was used to determine whether two criteria would be considered congruous. For example, LOEM's (Kay and Knaack, 2008) criterion of Meaningful Interactions and Lester's (2000) criterion of Don’t use animation as a substitute for interaction, while not identical, were deemed similar enough to be considered the same criterion. In total, 32 criteria were determined to appear in multiple instruments. This set of criteria was then discussed with an expert and considered along with the literature on the design of educational technology tools. As this instrument is intended to evaluate an app based on how it will enhance students' learning, those criteria that were unrelated to the learning of students (e.g. credits to authors are provided) were then eliminated. This process resulted in the sixteen criteria used in this instrument, which can be seen in Table 1. The entire instrument (the criteria, along with guidelines for assigning scores) can be found in Appendix K.
The criteria are grouped into four categories. An app is given a score from 1 to 5 for each of the criteria. A 5 would indicate the app meets the criteria very well, with any shortcomings being very minor and likely to not affect the student's learning. A 4 would indicate the app adequately meets the criteria most of the time, but with some flaws occasionally present or flaws that would only affect some users. A 3 would indicate that there is significant room for improvement or that there are a handful of flaws that have the potential to hinder students' learning. A 2 would indicate that the criterion is not well-met, and that there are significant flaws present that are likely to hinder students' learning. A 1 would indicate that the app does not meet the criterion at all, with major flaws, ones that could potentially prevent the app from being used as an effective learning tool.

The scores for each of the criteria within a category are averaged to produce a score for each category. These four category scores are then averaged to produce a single overall score for the app. The overall score is calculated in this fashion so as to not give extra weight to categories with more criteria within them. What follows is a discussion and justification of each of the criteria.
Table 1. The Criteria Used in the Evaluation Instrument

<table>
<thead>
<tr>
<th>Category</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interactions</td>
<td>Meaningfully interactive</td>
</tr>
<tr>
<td></td>
<td>Natural actions on objects</td>
</tr>
<tr>
<td></td>
<td>Visible consequences of actions</td>
</tr>
<tr>
<td></td>
<td>Range of actions not limited by technology</td>
</tr>
<tr>
<td></td>
<td>Necessitate thoughtful, deliberate actions</td>
</tr>
<tr>
<td></td>
<td>Cognitively faithful interactions</td>
</tr>
<tr>
<td>Quality of Content</td>
<td>Mathematically faithful</td>
</tr>
<tr>
<td></td>
<td>Cognitively demanding</td>
</tr>
<tr>
<td></td>
<td>High-quality presentation</td>
</tr>
<tr>
<td>Feedback and Support</td>
<td>Provides relevant and timely feedback</td>
</tr>
<tr>
<td></td>
<td>Provides scaffolding</td>
</tr>
<tr>
<td></td>
<td>Provides opportunities for reflection</td>
</tr>
<tr>
<td>Usability</td>
<td>Easy to use</td>
</tr>
<tr>
<td></td>
<td>Documentation is clear and easy to understand</td>
</tr>
<tr>
<td></td>
<td>Interface is clear, consistent, and follows technology standards</td>
</tr>
<tr>
<td></td>
<td>Usable by a variety of learners</td>
</tr>
</tbody>
</table>

The Interactions Category

Apps for learning mathematics should be meaningfully interactive. The app should consist of more than just a series of animations or static screens that the user simply watches and advances through (Lester, 2000). The reasons for this are twofold. First, allowing a student to manipulate on-screen objects has been shown to result in higher retention and increased problem-solving skills over students that watch non-interactive animations (Chan & Black, 2006). Second, these types of activities promote a student-centered, active learning environment, advocated by constructivist-learning theorists.
An app that receives a 5 for this criterion would require that the user be an active participant at all times, controlling what is happening on-screen. The user should have a choice of actions, and the on-screen objects should be transformed because of these actions. An app that would receive a 3 may be one in which some mathematical decisions are made without input from the user, or one in which the user spends a moderate amount of time as a passive observer. An app that receives a 1 would require little interaction from the user, particularly when it comes to the mathematics being performed.

The mathematical actions that are performed should be on objects (Lester, 2000) and should be natural – that is, when an action is performed, the object should behave as expected, either because of the properties of the mathematical entity represented, or because of standards adopted. Note that this not only depends on the action being performed, but also on how the action is performed. Tversky, Morrison, and Betrancourt (2002) define the congruence principle as “the structure and content of the external representation should correspond to the desired structure and content of the internal representation” (p. 249). In other words, the way in which an action is physically performed (external) should match how the user perceives the mathematical action that is happening (internal). For example, if one was to rotate an on-screen figure by performing a rotation motion with their fingers, this would be a congruent action, while tapping on the figure would not be. Segal (2011) showed that in a touch-interface, students perform better on tasks that require congruent actions than those that require incongruent actions.

For an app to receive a 5 for the natural actions on objects criterion, all (or nearly all) actions done to mathematical objects (e.g., expressions, variables) should be accomplished
by gestures with the object, and these gestures must be congruent to the mathematical action being performed. An app that receives a 3 may occasionally, in order to manipulate an object, require the user to press a button away from the object or select a menu option. Or some gestures may feel unnatural or be incongruent to the mathematical action being performed. An app that receives a 1 is likely to consist of mathematical objects that can not be touched or directly interacted with.

The consequences of the user's actions should be clearly visible. Perhaps this means momentarily highlighting an object before it changes, or emphasizing the link between the independent object that can be changed and the dependent object or objects that will change as a result. Meeting this criterion will help students better understand relationships (Underwood et al., 2005), a fundamental part of learning algebra.

For an app to receive a 5 for this criterion, the consequences of all or nearly all of the user's actions must be visible, and changes that happen are expected and noticeable. An app that receives a 3 may have some actions that happen instantaneously, without any animation, or has changes happen far away from the user's actions. An app that receives a 1 would often have changes that happen unexpectedly, or have multiple actions happening simultaneously, and it would not be clear which objects have changed and how.

The visibility of the consequences of one's actions contributes to the app's pedagogical fidelity, which Zbiek et al. (2007) define as “the extent to which teachers (as well as students) believe that a tool allows students to act mathematically in ways that correspond to the nature of mathematical learning that underlies a teachers practice” (p.
It “is about allowing students to ‘do’ mathematics and not be distracted or limited by technical features” (Bos, 2009, p. 522).

Thus, as another requirement to achieve pedagogical fidelity, the user's range of actions which can be performed on objects should not be limited by the technology. It is students' reflections on these mathematical actions that lead to them making predictions and testing these predictions, ultimately leading to a greater understanding of the mathematical concepts and objects being represented on-screen (Dick, 2008). Therefore, when actions that the user can perform are limited, the potential for students' learning is diminished.

An app which receives a 5 for this criterion must use the technology to its full potential, allowing the user to perform nearly any actions that they could with pen-and-paper, and possibly more. An app that receives a 3 may restrict some of the actions the user can perform or restricts when they can perform certain actions. The technology likely offers little advantage over pen-and-paper methods. In an app that receives a 1, the user may often be forced to perform certain mathematical actions. Little opportunity is afforded for making and testing predictions, and alternate solution paths that would otherwise be viable are not able to be traversed.

An effective mathematics app should necessitate thoughtful, deliberate actions, rather than random ones. If a student completes a task by randomly performing actions, he or she is unlikely to learn the intended mathematical concepts or improve their problem-solving skills.

In an app receiving a 5 for this criterion, it is unlikely that a student could progress without carefully considering his or her actions. For an app to receive a 3, a typical student may be able to progress using random actions, but more deliberate actions may be more
efficient. Or perhaps the set of possible solutions is relatively small (e.g., integers between -10 and 10) without any penalty for incorrect answers, so users could conceivably arrive at the correct solution by simply guessing, even though the intended solution path may require thoughtful, deliberate actions. An app receiving a 1 may be one that the user is expected to randomly guess or use trial-and-error without first considering appropriate guesses. Even if random guessing is not explicitly encouraged, it may be that most users will immediately find that random guessing is more efficient, and have no reason to not resort to random guessing.

Taken as a whole, the interactions afforded by the app should exemplify the app's cognitive fidelity. Cognitive fidelity refers to “how well the virtual tools reflect the user’s cognitive actions and possible choices while using the tool in the virtual environment” (Moyer-Packenham et al., 2008, p. 204), or more succinctly, “whether a concept is better understood when the object is acted on” (Bos, 2009, p. 522). In the “real world”, it is often acceptable for technological tools to arrive at “black box” solutions as efficiently as possible; however, in the classroom, learning apps should place a priority on illuminating mathematical thinking processes (Dick, 2008). For an example of how an algebra app might be cognitively unfaithful, consider the graph of the equation \( x^2 + y^2 = 1 \) (i.e., a circle of radius 1 centered on the origin). If the app allows the user to scale the horizontal axis independently from the vertical axis, the circle seems to turn into an ellipse, with some points seemingly closer to the center than others. Particularly for a lower-level algebra student, this can cause confusion and obstruct the mathematical properties of a circle.
For an app to receive a 5 for the *cognitively faithful interactions* criterion, the interactions afforded by the app must clearly convey the properties of the mathematical objects within it. The app should elucidate why and how mathematical procedures and methods work. An app that receives a 3 may fail to explain several mathematical elements within it. The user may be expected to blindly apply procedures without understanding why the procedures work. An app that receives a 1 is likely to be confusing for users. Feedback (implicit or explicit) may be misleading, and while the content may be mathematically accurate, the user is not likely to gain much mathematical knowledge from using the app.

*The Quality of Content Category*

The mathematics in any app should be accurate; the app should possess *mathematical fidelity*. Mathematical fidelity is defined by Moyer-Packenham et al. (2008) as “the degree to which the mathematical object is faithful to the underlying mathematical properties of that object in the virtual environment” (p. 204). In the algebra classroom, these “mathematical objects” typically include such things as variables, real numbers, operations, equality, functions, and graphs. While mathematical accuracy may seem like an obvious goal to include in any learning app, it is not always easy to accomplish. Sometimes, it is the limits of the technology that are the obstacle: numerical precision, a pixelated display, the modeling of continuous structures with discrete ones. Other times, conscious design decisions are made to sacrifice mathematical fidelity for ease of use or to match user expectations. For example, Dick (2008) points to the interpretation of the user-entered expression sin2x. Many apps will take this to mean sin(2x), giving implicit multiplication a higher priority than application of a function. While this may be a convenience to the user, it introduces the unusual situation in
which the implicit multiplication in \( \sin 2x \) is treated differently than the explicit multiplication in \( \sin 2 \times x \), where the application of the sine function would happen before the explicit multiplication (p. 335).

For an app to receive a 5 for mathematical fidelity, all mathematical content must be accurate. An app that receives a 3 contains mathematical errors, but ones that might not be encountered by all users, or are errors of insufficient or incomplete information, rather than untruths. An app that receives a 1 has major mathematical errors.

Apps for learning mathematics should be cognitively demanding (Bowers, Bezuk, & Aguilar, 2011; Dick & Hollebrands, 2011; Underwood et al., 2005). That is, they should require higher-level thinking rather than simple memorization or application of procedures. Note that perceived task difficulty is not necessarily correlated with cognitive demand. In algebra, for example, memorizing and using the quadratic formula may be difficult for some students, but an app that simply requires the application of the quadratic formula to several equations in the form \( Ax^2 + Bx + C = 0 \) is not cognitively demanding (Bowers et al., 2011).

For an app to receive a 5 for this criterion, it must evoke critical thinking from the user. The user should be expected to synthesize knowledge previously learned – either from earlier stages in the app, or from information the user is expected to know before using the app – with new information in order to solve the task at hand in non-procedural way. An app that receives a 3 for this criterion may have some tasks that are repetitive, or may give too much guidance or direction to the user. An app that receives a 1 is comprised of trivial or menial tasks.
The quality of the presentation should be high (Kay & Knaack, 2008a; Nesbit et al., 2009; Underwood et al., 2005). This criterion is not the same as an attractiveness or an aesthetics criterion – the quality of the presentation should be judged based on how it enhances learning, not as one would judge the beauty of a piece of art. Multimedia in the app, including text, audio, graphics, video, and animations, should be clear and concise. Multiple types of multimedia should be incorporated in ways that enhance learning; in particular, the app should not be overloaded with needless or wordy text. Text and other media should be free of distracting typographical or technical errors.

For an app to receive a 5 for this criterion, all components of the presentation should contribute to (or at the very least, not impede) student learning. Sounds, colors, and animations could be used as various indicators to highlight properties or changes. An app that receives a 3 may have some distracting or misleading components, or some graphics, text or audio may be unclear. An app that receives a 1 has very little beneficial graphics or sound. Text or pictures may be unreadable, or there are unnecessarily wordy and lengthy blocks of text.

The Feedback and Support Category

The mathematics app should provide relevant and timely feedback (Kay & Knaack, 2008a; Le@rning Federation, 2003; MERLOT, 2013; Vargo et al., 2003). This feedback can take several forms. It can give the status of the problem being solved or of the mathematical objects being examined. It can be conceptual feedback, asking the user questions that cause them to reconsider their perceptions of the objects. It can be corrective feedback, letting the user know a mistake was made and guiding the user to correct the mistake. In any case, the
feedback should provide formative assessment – that is, assessment for learning, rather than simply assessment of learning (Bokhove & Drijvers, 2010).

For an app to receive a 5 for this criterion, formative feedback must be provided at various stages throughout the task. The feedback provided must help the student progress, both in the task and in their understanding of the mathematics involved. An app that receives a 3 may provide frequent feedback, but only of the corrective type, or relevant feedback may be present during some parts of the app, but lacking in others. An app that receives a 1 provides little to no feedback to the user.

The mathematics app should provide scaffolding (Haughey & Muirhead, 2005; Learning Federation, 2003; MERLOT, 2013). The scaffolding should ensure that the user never feels like they are conceptually “stuck” or “lost” when it comes to advancing in the task. The scaffolding can be explicit, in the form of hints or guiding questions, or it can be implicit, by gradually increasing difficulty and building upon concepts.

An app that receives a 5 for this criterion provides both implicit and explicit scaffolding. Steps are taken to ensure that the user is always able to progress, while at the same time, building their knowledge of the mathematical topics being examined. An app that receives a 3 may provide implicit scaffolding, but lack on-demand or as-needed assistance. Or if real-time assistance is provided, the app may not be requiring the user to build upon concepts learned earlier in the app. In an app that receives a 1 for this criterion, no recourse is provided for a user that becomes stuck, and there is no sensible progression of tasks.

The mathematics app should provide opportunities for reflection (Clements, 2000; Dick & Hollebrands, 2011; Haughey & Muirhead, 2005; Underwood et al., 2005). This could
take the form of a history of actions, or sense-making questions near the end of the task, or simply deliberate pauses throughout the task. Discussion could also be included regarding how the mathematics relates to the real world or to other fields of study.

For an app to receive a 5 for this criterion, it must actively promote and stimulate reflection. An app that receives a 3 may provide opportunities for reflections, but ones that are likely to be skipped or ignored by the user, or ones that simply do a poor job of promoting cogitation. An app that receives a 1 moves instantly from one topic or problem to the next, perhaps even encouraging the user to progress as quickly as possible, and asks the user no thought-provoking questions.

*The Usability Category*

The mathematics app should be *easy to use* (Dick & Hollebrands, 2011; Kay & Knaack, 2008a; MERLOT, 2013). Most students should be able to successfully use the app without a lot of documentation, technical skills, or instruction.

For an app to receive a 5 for this criterion, it must be “pick-up-and-play”. That is, the typical user should be successfully progressing through the tasks moments after starting the app. An app that receives a 3 has some features that are not clearly evident without consulting documentation, or the goal of the task is not immediately clear. An app that receives a 1 requires the user to frequently refer to the documentation for commands or actions, and may require significant technical skills to progress in the task.

When documentation or help files are present or necessary, *they should be clear and easy to understand* (CLOE, 2004; Haughey & Muirhead, 2005; Kay & Knaack, 2008a).
An app that receives a 5 for this criterion has clear help files with pictures, animations, or examples. Help files are easy to navigate, and specific topics are easy to find. An app that receives a 3 may have help files that are somewhat wordy or unclear. Or the help files themselves are clear, but finding help on a specific topic is difficult, or accessing the help files requires the user to abandon the current task. An app that receives a 1 has little to no documentation or help files, despite the need for them. If the app does have documentation, it is often incorrect, hard to understand, or otherwise not helpful. Note that the amount of help files does not necessarily correlate with this score – an app that does not have a need for many help files should be not dispraised for having few.

The user-interface of the app should be clear, consistent, and follow technology standards. A clear interface means that text should be large enough to be readable, buttons or other objects that can be clicked are clearly identified, and unnecessary information or dialogues are kept to a minimum. A consistent interface means that all or most pages have a similar look and feel and that text and visual components are always in the same location. Following technology standards means, for instance, that the user is expected to tap an object to select it or that icons for commonly used functions (save, pan, zoom) are instantly recognizable by the typical iPad or computer user.

For an app to receive a 5 for this criterion, the interface must be clear and consistent throughout the entirety of the app, and where applicable, technology standards should be followed. An app that receives a 3 may have some buttons or interactive objects that are not identified as such, or the placement of questions, buttons, or other objects may vary somewhat throughout the app, or buttons may have non-standard icons or otherwise not
clearly identify their purpose. An app that receives a 1 may have text that is too small to be readable, or buttons that are too small to be easily clicked on. Menus, text, or other objects may extend past the edges of the screen or overlap each other, causing them to be hard or impossible to read or interact with. Buttons may be unlabeled.

The app should be *usable by a variety of learners* (Bokhove & Drijvers, 2010; CLOE, 2004; Haughey & Muirhead, 2005; Le@rning Federation, 2003; Underwood et al., 2005). Allowing multiple entry points, multiple approaches, and multiple solution strategies enables students of varying abilities and learning styles to express their own mathematical ideas, and engage with and complete the task. Multiple difficulty levels is another approach that can tailor the app to the learner's ability. The app can be effective for an even wider range of learners if multiple languages are supported, or if the app complies with accessibility standards for students with disabilities.

For an app to receive a 5 for this criterion, it must allow multiple strategies and be accessible to students of varying learning styles and mathematical abilities. An app that receives a 3 may have problems of multiple difficulty levels, but problems that do not allow multiple solution paths. If the app contains problems that allow multiple solution paths, the problems are not tailored to the ability level of the student. An app that receives a 1 shoehorns all users into identical problems with one solutions path. Although multiple-language support or meeting accessibility standards are not required to receive a 5, they can improve the score received.
Remarks

Some remarks about the instrument are warranted. First, some things that are not included in the instrument will be addressed.

There is no motivation category or criterion. While motivation appeared in some of the instruments reviewed in Chapter 2, or in their supplementary materials, only two (Nesbit et al., 2009; Underwood et al., 2005) provided any concrete examples or ways that a learning object might motivate a student. It is arguable, however, that an app that meets the other criteria in this instrument would already be motivating to students. Take, for example, an app that is easy to use, has a high-quality presentation, is cognitively challenging, provides timely feedback and scaffolding, and is perceived by a student to have a high learning value. The app is likely to be enjoyed by the student and he/she may be motivated to use it simply for the fact that the student thinks it will help him/her pass the course. Indeed, most of the motivating qualities found in Nesbit et al. (2009) and Underwood et al. (2005) (e.g., high-quality media, interactivity, quality feedback, providing early easy questions or activities) are already encompassed by other criteria in this instrument.

Aesthetics also does not appear in this instrument. As noted earlier, there is a difference between aesthetics and the quality of presentation as it relates to learning. Some of the instruments that mention the attractiveness of the app do not differentiate between the two, and the criteria they provide are more closely related to a criterion already presented in this instrument than to a potential aesthetics criterion. For example, MERLOT's (2013) criterion of “Does the material present information in ways that would be attractive to students?” is in its Ease of Use category, and Bokhove and Drijvers's (2010) attribute of “The
structure and attractiveness of the tool’s interface” is focused on how the user interacts with the learning object. LORI 1.3 (Vargo et al., 2003) included aesthetics as one of its ten criteria, but later versions (Nesbit et al., 2004; Nesbit et al., 2009) no longer included aesthetics, and instead incorporated aesthetic quality into its presentation design criterion. While flashy gimmicks such as glitzy banners and irrelevant audio tracks may initially grab a user's attention, their use in learning objects rarely leads to substantive motivational benefits, and soon become distractions that can interfere with learning (Leacock & Nesbit, 2007).

It should also be pointed out that unlike some evaluation instruments (e.g., Kay & Knaack, 2008a), it is not intended that different categories be separate constructs, independent of each other. For example, if an app is not cognitively demanding (quality of content category), it is unlikely to require thoughtful, deliberate actions (interactions category). If the interface is confusing (usability category), it is less likely that the consequences of the user's actions will be visible (interactions category).

**Choosing Apps to Evaluate**

In order to evaluate the instrument itself and to discover any difficulties that might arise when applying the instrument, a sample of apps was chosen to apply the instrument to.

A multitude of apps were located via a variety of methods, including personal recommendation, web searches, and keyword searches in the App Store. Based on product descriptions and screenshots, the selection was filtered to include only those apps which were related to algebra and were educational (e.g., a quadratic equation solver that simply outputs the roots of an inputted function was not included). This resulted in approximately 25 apps that were downloaded for further inspection. After spending a few minutes with each app,
this selection was again filtered to include only those apps that were interactive (e.g., digital versions of textbooks were not included) and intended to help users solve algebraic equations. This process resulted in seven apps, seen in Table 2, which served as the initial sample to be discussed and evaluated using the evaluation instrument.

**Applying the Instrument & Data Analysis**

Using the evaluation instrument, I evaluated each app, assigning scores from one to five for each of the 16 criteria. These scores were then discussed with the chair of the advisory committee, and adjusted if major concerns were raised. In some cases, this discussion led to some of the criteria being reevaluated or reworded to make them clearer, until an agreed upon score could be reached.
Table 2. The apps selected for evaluation

<table>
<thead>
<tr>
<th>App</th>
<th>Developer</th>
<th>Release Date</th>
<th>Version Reviewed¹</th>
<th>iTunes link</th>
<th>Price</th>
<th>App Store Rating²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numerosity: Play with Math! (Ch. 5)</td>
<td>Thoughtbox</td>
<td>Nov. 6, 2012</td>
<td>1.1.1</td>
<td><a href="https://itunes.apple.com/ie/app/numerosity-play-with-math!/id554637363%C2%B3">https://itunes.apple.com/ie/app/numerosity-play-with-math!/id554637363³</a></td>
<td>$1.99⁴</td>
<td>4.0 (entire app)</td>
</tr>
</tbody>
</table>

¹ All apps were reviewed using the latest version available on the App Store as of April 20, 2013
² As of April 20, 2013. Per App Store policy, average ratings are rounded to the nearest 0.5, and are for the most current version only, unless less than five ratings are available for the most current version.
³ No longer available
⁴ To unlock Chapter 5. Downloading the app itself is free.
Once this process was complete, a cross-app analysis was performed to examine commonalities among the seven apps. Mean scores for each criteria and each category were used to determine in which areas these apps generally performed well in, and in which areas the apps were commonly lacking.

For each of the 16 criteria, the Pearson product-moment correlation coefficient was calculated for the ratings from the App Store versus scores for each of the seven apps in that criteria. This was done to determine if any of the criteria were good predictors of App Store ratings. The same analysis was also done for the four category scores, as well as the overall app scores.

To help assess the instrument itself, a fellow master's student in Mathematics Education was presented with the instrument (similar to what appears in Appendix K), and asked to evaluate two of the seven apps – a random selection from the three apps receiving the highest overall ratings in the initial review, and a random selection from the three apps receiving the lowest overall ratings. Any questions that the student had about definitions of terms or layout of the instrument were answered; otherwise, no training on any of the criteria or on employing the instrument was provided. From the scores assigned during this evaluation, the inter-rater reliability was calculated to determine agreement between the two raters and reliability of the instrument. Additionally, the student was asked to comment on the ease of using the instrument, and on any individual criteria that may have been difficult to assess.
RESULTS OF APP EVALUATIONS

Seven iPad apps for learning about algebraic expressions and equations were initially evaluated using the instrument. The scores resulting from this evaluation are shown in Table 3. A short description of each app follows, along with a justification for each of those scores.

Hands-On Equations for the iPad

Hands-On Equations is a series of apps based on the physical manipulatives of the same name\textsuperscript{5}. The app, like the manipulatives, can, according to the developer, be used by students from the age of eight, or by older students struggling with algebra.

The user is presented with a linear algebraic equation. Using two different colors of pawns (representing $x$ and $-x$) and two different colors of numbered cubes (representing positive and negative integers), the user builds each side of the equation on the two sides of a scale, as seen in Figure 1.

\begin{flushleft}
\textsuperscript{5} See http://www.borenson.com/ for details regarding the Hands-on Equations physical manipulatives.
\end{flushleft}
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Meaningfully Interactive</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>4.6</td>
<td>0.79</td>
</tr>
<tr>
<td>Natural actions on objects</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3.6</td>
<td>1.40</td>
</tr>
<tr>
<td>Visible consequences of actions</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>3.9</td>
<td>1.46</td>
</tr>
<tr>
<td>Range of actions not limited by technology</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2.6</td>
<td>0.79</td>
</tr>
<tr>
<td>Necessitate thoughtful, deliberate actions</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3.4</td>
<td>1.13</td>
</tr>
<tr>
<td>Cognitively faithful interactions</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2.7</td>
<td>0.76</td>
</tr>
<tr>
<td><strong>Interactions category overall score</strong></td>
<td><strong>3.8</strong></td>
<td><strong>4.2</strong></td>
<td><strong>4.2</strong></td>
<td><strong>2.7</strong></td>
<td><strong>3.8</strong></td>
<td><strong>2.7</strong></td>
<td><strong>2.8</strong></td>
<td><strong>3.45</strong></td>
<td><strong>0.70</strong></td>
</tr>
<tr>
<td>Mathematically faithful</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3.4</td>
<td>1.27</td>
</tr>
<tr>
<td>Cognitively demanding</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>3.1</td>
<td>0.90</td>
</tr>
<tr>
<td>High-quality presentation</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>3.7</td>
<td>1.38</td>
</tr>
<tr>
<td><strong>Quality of content category overall score</strong></td>
<td><strong>4.3</strong></td>
<td><strong>4.0</strong></td>
<td><strong>4.0</strong></td>
<td><strong>1.7</strong></td>
<td><strong>4.3</strong></td>
<td><strong>2.3</strong></td>
<td><strong>3.3</strong></td>
<td><strong>3.43</strong></td>
<td><strong>0.97</strong></td>
</tr>
<tr>
<td>Provides relevant and timely feedback</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>3.0</td>
<td>1.41</td>
</tr>
<tr>
<td>Provides scaffolding</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3.4</td>
<td>0.53</td>
</tr>
<tr>
<td>Provides opportunities for reflection</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2.4</td>
<td>0.98</td>
</tr>
<tr>
<td><strong>Feedback and support category overall score</strong></td>
<td><strong>2.3</strong></td>
<td><strong>3.7</strong></td>
<td><strong>2.7</strong></td>
<td><strong>2.0</strong></td>
<td><strong>4.3</strong></td>
<td><strong>2.7</strong></td>
<td><strong>3.0</strong></td>
<td><strong>2.95</strong></td>
<td><strong>0.8</strong></td>
</tr>
<tr>
<td>Easy to use</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>3.9</td>
<td>1.21</td>
</tr>
<tr>
<td>------------------------------</td>
<td>--------------------</td>
<td>---------------</td>
<td>-------------------</td>
<td>-----------------------------</td>
<td>------------</td>
<td>---------</td>
<td>--------------------------------</td>
<td>---------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Documentation is clear and easy to understand</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>3.6</td>
<td>1.13</td>
</tr>
<tr>
<td>Interface is clear, consistent, and follows technology standards</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4.0</td>
<td>1.00</td>
</tr>
<tr>
<td>Usable by a variety of learners</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4.1</td>
<td>1.07</td>
</tr>
<tr>
<td>Usability category overall score</td>
<td><strong>4.0</strong></td>
<td><strong>4.8</strong></td>
<td><strong>4.5</strong></td>
<td><strong>4.0</strong></td>
<td><strong>3.3</strong></td>
<td><strong>3.5</strong></td>
<td><strong>3.3</strong></td>
<td><strong>3.89</strong></td>
<td><strong>0.59</strong></td>
</tr>
<tr>
<td>Total overall app score</td>
<td><strong>3.63</strong></td>
<td><strong>4.15</strong></td>
<td><strong>3.83</strong></td>
<td><strong>2.58</strong></td>
<td><strong>3.94</strong></td>
<td><strong>2.79</strong></td>
<td><strong>3.1</strong></td>
<td><strong>3.43</strong></td>
<td><strong>0.61</strong></td>
</tr>
</tbody>
</table>
Using procedures that the app teaches, the user manipulates the objects until pawns of a single color are on one side of the scale, and blocks of a single color are on the other side. At that point, the user is expected to use deduction to figure out the value of a pawn.

*Meaningfully Interactive: 4*

Each lesson begins with a non-interactive video, usually two to three minutes in length. These videos must be watched before the practice problems become available. Outside of the videos, however, the student is in control, and can solve problems at his or her
own pace. The student's time is spent building equations by placing objects onto the scale, and then manipulating the objects to isolate the pawns.

*Natural actions on objects: 5*

All actions are performed directly on the pawns and blocks. Objects are moved by dragging, and adding and removing objects from the balance is achieved by simply dragging them onto or off the balance. The action of adding an object to each side of the scale is congruent with the mathematical action of adding the same amount to each side of the equation.

*Visible Consequences of Actions: 5*

It is clear whether pawns and blocks are on the balance or not; objects moved to the balance are automatically adjusted so that they sit directly on the balance, rather than floating nearby. When using the auto-check feature to sum the objects on each side of the balance, each object pulses when it is being added, and the intermediate sum is displayed, letting the user follow along as each object is added, as seen in Figure 2.

*Range of actions not limited by technology: 3*

The virtual objects – pawns, blocks, and the balance – in this app emulate the Hands-on Equations physical manipulatives, which in turn were developed to represent mathematical objects – variables, integers, and equality. The app faithfully emulates the physical manipulatives; however, this means that any deficiencies in the representation of the mathematical objects as the physical manipulates are also present in this app. For example, the red 6 block in this app represents the integer 6. We can not, however, subtract 2 from this, leaving a 4 block, without first removing the 6 block from the balance, placing a 2 block and
a 4 block on the balance, and then removing the 2 block. Likewise, we can not split a 6 block into the sum of multiple blocks, like one could do with an integer. These are common and useful actions when solving problems like the ones presented in this app. For a more striking example, consider subtracting 5 from 2. Starting with a red 2, we would have to add a zero in the form of a red 5 and a green 5, then take away the red 5 to perform our subtraction, then remove the green 5 and replace it with a green 3 and a green 2, and then take away a zero by removing the green 2 and the red 2, leaving the green 3.

*Figure 2. The equation has been set up on the scale.*
Necessitate thoughtful, deliberate, actions: 3

The app does not require the use of the balance. With all solutions being a non-zero integer, small in magnitude, a user could conceivably randomly try numbers until the correct solution was found. The app also asks the student to check their solution by evaluating each side of the equation when the variable is replaced with the potential solution, though again, the use of the balance is not required, and guessing can suffice. In addition, an auto-check feature is included that will perform this operation for the user, as long as he or she has set up the equation on the balance. While this is presumably included as a convenience to the user, the use of it removes the need for thoughtful action.

Cognitively faithful interactions: 3

The central object in this app is the balance, representing equality. In learning and performing actions on this balance such as removing or adding equal quantities to both sides, the user should come to understand the relational properties of equality. The quantity $4x$ is represented by a set of four objects, while the quantity $4$ is represented by a single object. This means that simplifying an expression like $4x - 3x$ requires different types of actions than simplifying $4 - 3$. In addition, a student could conceivably complete the entire app and still not be able to quickly simplify expressions like $x + 2 - 3$, or $x + 2 + (-3)$, or even know that these two expressions result in the same term, since they would require significantly different actions in this app, as seen in Figure 3.
| 1. Begin with blue pawn and red 2 | 1. Begin with blue pawn, red 2, and green 3 |
| 2. Add green 3 and red 3 | 2. Replace green 3 with green 2 and green 1 |
| 3. Remove red 3 | 3. Remove green 2 and red 2 |
| 4. Replace green 3 with green 2 and green 1 | 4. End with blue pawn and green 1 |
| 5. Remove green 2 and red 2 | |
| 6. End with blue pawn and green 1 | |

*Figure 3. Steps required to set up and simplify two similar expressions*

**Mathematically faithful:** 5

No errors were observed in the ways that the mathematical objects were represented.

**Cognitively Demanding:** 4

Each lesson adds a slightly different element than the last, requiring that users synthesize old knowledge with new. However, within each lesson, the methods taught are largely procedural, although the user must still decide which procedures are appropriate to use at the given moment.

**High-quality Presentation:** 4

Videos are clear and concise, objects are easy to identify, colors are used (but not consistently) to signify negative and positive objects.

**Provides relevant and timely feedback:** 2

While performing the problems, the app gives almost no feedback. The app does not indicate when the problem has been set up on the scales correctly. When the user adds or removes objects, the app does not indicate if the scales are still balanced. Feedback given at the end of each problem is corrective only (e.g., “x is not -4” or “try the check again”)

---
<table>
<thead>
<tr>
<th>$x + 2 - 3$</th>
<th>$x + 2 + (-3)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Begin with blue pawn and red 2</td>
<td>1. Begin with blue pawn, red 2, and green 3</td>
</tr>
<tr>
<td>2. Add green 3 and red 3</td>
<td>2. Replace green 3 with green 2 and green 1</td>
</tr>
<tr>
<td>3. Remove red 3</td>
<td>3. Remove green 2 and red 2</td>
</tr>
<tr>
<td>4. Replace green 3 with green 2 and green 1</td>
<td>4. End with blue pawn and green 1</td>
</tr>
<tr>
<td>5. Remove green 2 and red 2</td>
<td></td>
</tr>
<tr>
<td>6. End with blue pawn and green 1</td>
<td></td>
</tr>
</tbody>
</table>
"Provides scaffolding: 3

The app starts with simpler lessons, and gradually builds into more complicated equations to solve. However, within each problem, the user is almost completely on their own from beginning to end, with the only available assistance being watching the introduction video again.

"Provides opportunities for reflection: 2

There is no history of actions, there are no sense-making questions, and the videos provide little outside of demonstrating procedures. There is a requirement that the user check his or her answer, but that too is presented as another procedure, rather than an opportunity to examine what the “answer” means in the context of an equation.

"Easy to use: 3

If the user is already familiar with the physical version of Hands-on Equations, the app is very intuitive, and should be able to be used with little to no instruction. If the user is not familiar with the physical version, however, it is hard to imagine that the typical student would be able to use the app to solve problems without first watching the instructional videos.

"Documentation is clear and easy to understand: 3

The entirety of the documentation is in the form of videos. The videos themselves are clear and easy to understand. However, there is no way to find help on a specific topic, unless you happen to know what lesson (identified only by numbers) the topic was introduced in, and even then, the videos must either be visually searched or watched in their entirety to find the information pertaining to the topic."
Interface is clear, consistent, and follows technology standards: 5

Problems are consistently presented, with objects and text in the same location each time. The interface, both in the menus and during a problem, is clear of superfluous text or buttons. Menus, buttons, and scroll-wheels behave as expected.

Usable by a variety of learners: 5

There are several problems of varying difficulty. For most problems, there are several viable strategies a student can use. These different strategies are presented in the videos, and students are told that they can use whichever strategy works best for them.

Algebra Touch

Algebra Touch was created by an individual software developer, with support from the University of Indiana and the University of Richmond. Funding for research (discussed briefly in Chapter 5) on the use of the app was provided by the U.S. Department of Education. Versions of the app are also available for Mac OSX and Windows 8.

In Algebra Touch, the user can manipulate algebraic expressions and equations by rearranging terms, performing operations, factoring, and a number of other actions. The user can progress through a series of interactive lessons, from early topics such as addition to topics such as factoring and advanced lessons that incorporate several topics. Each lesson is accompanied by a Practice Mode, which randomly-generates problems related to the lesson, or the user can create and solve their own problems. A typical lesson is shown in Figure 4.
Meaningfully interactive: 5

When solving a problem, the user is in control and constantly interacting with the expression or equation on screen. The user can perform various actions to alter the expressions or equations, including rearranging terms or moving them to the other side of the equation, performing arithmetic operations to simplify, and canceling out common factors in the numerator and denominator. Even in the Explain Mode, all examples are interactive, and the user follows along with the explanations by interacting with the on-screen expression.

Figure 4. A typical lesson
Natural actions on objects: 5

Actions are performed directly on objects and are congruent with the mathematical action taking place. For example, terms are rearranged by dragging them, common factors in the numerator and denominator are “canceled out” by drawing a line through them, and expressions are factored by “pulling out” the common factor. The result of this is that the objects being interacted with are likely to behave in the way that the user expects.

Visible consequences of actions: 5

Equations are constantly updated as terms are being dragged, so it is easy for the user to see what the result of their action will be before completing it. When operations are performed, the operands clearly fade out to be replaced by the result. When invalid operations are attempted (e.g., attempting to simplify the addition of an integer and a variable, or violating order of operations), the invalid operation and operands are visibly and audibly identified. When distributing multiplication over addition, the multiplicand is clearly shown to multiply both addends, as seen in Figure 5.
Range of actions not limited by technology: 3

Exponents of variables or integers are not supported. In other words, the app will not let you simplify $x^x$, and most problems with variables of a higher degree than 1 (e.g., $x^x = 9$) are not solvable. The user can drag operands to the other side of an equation, but only in some cases. For example, when solving $\frac{5}{3} = 5x$, one can drag the 3 in $\frac{5}{3}$ to the other side, but not the 5. Both the x and the 5 in $x^5 = 2$ can be dragged, but only the x in
\[ x \times \left( \frac{5}{3} \right) = 2 \]
can be dragged. One can not split fractions such as 
\[ \frac{3x + 8}{3} \]
into 
\[ \frac{3x}{3} \frac{8}{3} \]
or 
\[ \frac{x \times 4}{3} \]

into 
\[ x \times \left( \frac{4}{3} \right) \].

**Necessitate thoughtful, deliberate actions: 3**

Mathematically incorrect actions are not allowed, so it is conceivable that a student could solve problems by randomly moving terms around and tapping on operations. Thoughtful actions, however, are likely to be more efficient. Simplifying fractions can be performed by choosing appropriate factors for the numerator and denominator, but it can just as easily be done by blindly drawing a line through the fraction and letting the app appropriately factor the numerator and denominator automatically.

**Cognitively faithful interactions: 4**

The properties of equality are somewhat obscured by letting students drag terms to the other side of the equals sign, rather than performing the same operation to both sides. Arithmetic and algebraic operations and their properties (e.g., simplifying, distribution, factoring), however, are demonstrated well. Through the interactive lessons, the user is simultaneously being taught about the mathematics involved and experiencing it happen first-hand.
Mathematically faithful: 3

The app will let you divide both sides of an equation by an expression that can evaluate to zero, possibly causing solutions to be lost, as seen in Figure 6. This can cause problems to be considered “solved”, even though another solution was initially possible. The app will simplify zero divided by zero to zero. These errors are likely not to be encountered by the typical user, however, as they would be found only in a small number of custom-created problems.

Figure 6: A solution gets lost when dividing by $(x + 2)$
Cognitively demanding: 4

As the user progresses through the lessons, they are required to synthesize the knowledge learned in previous lessons with the newly introduced information. In the practice mode, though no two problems are identical, after several problems, the actions required to solve them could become somewhat procedural.

High-quality presentation: 5

Lessons are clear and concise. No superfluous text or buttons. Text is free of typographical and content errors. Colors are used to indicate what factors are involved when factoring or canceling out factors. Colors are also used to differentiate between objects that are part of the equation and those that are not. Animations are used to indicate when a expression is being replaced by a simpler one, or when a multiplicand is being distributed across addition. Audible and visible notifications are used to indicate invalid actions.

Provides relevant and timely feedback: 5

The app clearly indicates when an expression is fully simplified or an equation is solved. Invalid actions are immediately identified, both visually and audibly, and parts of the equation vibrate to indicate why the action was invalid. During the explain mode, feedback is given at multiple steps throughout the problem.

Provides scaffolding: 4

Actions and operations are introduced one at a time. Problems gradually increase in complexity. During explain mode, users are initially given step-by-step instructions that gradually abate until the users are solving problems on their own. Within each problem in
practice mode, however, students must complete the problem from beginning to end without any assistance.

Provides opportunities for reflection: 2

No history of actions. No sense-making questions. An undo button is present, but there is no indication that it should be used for review rather than to correct mistakes.

Easy to use: 4

If a student is familiar with the rules of algebra, they will likely be able to use the app successfully after a few minutes of experimentation.

Documentation is clear and easy to understand: 5

Each lesson is clear and is dynamically linked to an interactive example problem. Lessons are labeled with their contents (factoring, division, etc.), making content easy to find.

Interface is clear, consistent, and follows technology standards: 5

Problems are consistently presented. Gestures for zooming and panning are standard for iOS apps. Buttons are easily identified. Instructions, menus, and other icons fit cleanly on the screen.

Usable by a variety of learners: 5

Multiple strategies can be used to solve problems. Problems are of various complexities. Users can start at different lessons, depending on experience. Lists of factors are provided and arithmetic is performed by the app, so users with lower mathematical ability should still be able to succeed.
**DragonBox+ Algebra**

DragonBox+ Algebra touts itself as “the game that secretly teaches you algebra”. It was developed by Dr. Patrick Marchal, a Ph.D. in cognitive science, and Jean-Baptiste Huynh, a high school teacher. It has won various awards, including Best Children App and App of the Year at GULLTasten 2012, and Best Educational Game at Fun & Serious Game Festival 2012. DragonBox+ Algebra is targeted for children aged 5 to 12.

The app consists of 200 levels. In each level, there is a playing board with two sides. On this board are several illustrated cards, including “the box”. The object of each level is to isolate the box on one side of the board. Whirlpool cards can be removed by tapping on them; other cards can be combined with their “night” version to create a whirlpool card. An early level is shown in Figure 7.

*Figure 7. An early level*
As the levels progress, the presentation of the board and the cards slowly begin to change, and the user gains additional actions that he or she can perform. Later levels add a “deck” of cards from which you can select a card and place a copy of the card on both sides of the board. The “box” is replaced with a card with an $x$ on it. Whirlpool cards are replaced by cards displaying a zero. Letters start to replace the illustrations on other cards. Some of these changes can be seen in Figure 8.

*Figure 8. A few levels later. Several of the cards have changed*
In later levels, some cards are connected horizontally to other cards, signifying multiplication. Card can appear under other cards, signifying division. An equals sign appears in between the two sides of the board. Terms are joined by addition signs. Numbers start to appear on cards. “Night” versions of cards are replaced with cards with negative signs. Eventually, the user is solving what looks very much like an algebraic equation, as seen in Figure 9.

Figure 9. A level from late in the game
Meaningfully interactive: 5

The app consists almost solely of the user interacting with the cards. The user is constantly altering what appears on-screen by combining, adding, and eliminating cards.

Natural actions on objects: 5

All actions performed are directly on the cards, which represent the numbers and variables in an equation. Unlike most other apps of this type, the gestures and physical actions are intended to be learned before the user learns about the mathematical actions behind them. Since the user is learning the mathematical concepts as a result of performing the physical actions, there is a sort of forced congruency between the two.

Visible consequences of actions: 5

When cards are dragged, nearby potential positions are highlighted. When cards disappear, the disappearing animation lasts several frames. Changes always happen at the location the gesture is occurring. For these reasons, it is clear what is about to change, as well as what has just changed.

Range of actions not limited by technology: 2

Only a small subset of algebraic and arithmetic actions are allowed – integers can not be multiplied, subtraction is not present at all, etc. For the actions that are allowed, they are only allowed to be performed with certain integers or variables, or at certain times. This is certainly by design, however.

Necessitate thoughtful, deliberate actions: 5

Except for very early levels, it is unlikely that random actions would result in completing levels. The user must continually be aware of how the cards interact in order to
decide on the appropriate next move. To acquire three “stars” on a level, the user must solve the equation in the fewest amount of moves, and the expression opposite the variable must be fully simplified, which would certainly require carefully calculated actions.

*Cognitively faithful interactions: 3*

The properties of addition, inverses, multiplication, and equality are never explicitly explained. While leaving users to figure these things out on their own is a design choice – and by itself, not necessarily a poor one – their knowledge of these mathematical elements will likely be incomplete after using this app. When the user divides one term by a variable or an integer, the app forces the user to divide all terms by the same variable or integer before he or she can continue, but it is never explained why or even expected that the student knows why. Combining elements vertically creates a 1, but combining them horizontally creates a 0, which is also not explained. The app may help students apply procedures to solve equations, but will likely not help them better understand the mathematical objects within them.

*Mathematically faithful: 4*

In general, content is accurate. However, on several problems, certain mathematical actions are not allowed, presumably to make the problem easier to solve. But this can give the incorrect impression that, for example, one can sometimes multiply both sides of an equation by the same number, but not divide. Additionally, one of the problems given is

\[ x \times x = \frac{x}{3} \]  

The user is expected to divide by sides of the equation by x, to get the solution
\[ x = \frac{1}{3} \]. It is, in fact, not possible to get the other solution of \( x = 0 \), since the app only allows certain actions on this problem.

**Cognitively demanding: 3**

Most problems present a slightly different element than the problem before it (e.g., the box is replaced with an \( x \)). This means that users must synthesize the knowledge gained from previous problems, and make a hypothesis about how this new element connects to the previous ones. Other than figuring out the new element, however, problems are not significantly different enough that solving them does not become somewhat procedural.

Forcing users to complete certain actions before continuing (e.g., dividing all other terms by an integer after dividing one) means that users do not have to consider what the next appropriate step is. An example of this can be seen in Figure 10.

*Figure 10. After multiplying one term by \(-f\), the app indicates where else the \(-f\) card must go*
High-quality presentation: 5

The use of colors and the high quality of the illustrations on cards makes “opposite” cards easy to identify, for all types of cards. Sparkles are used to identify the “special” card (the box or the \( x \)). Animations and sounds are used to indicate changes in the equation or invalid actions. The different parts of the board are easily identified. Presentation is free of distracting text or other elements.

Provides relevant and timely feedback: 3

Invalid actions are visibly and audibly identified (although some otherwise valid actions are not allowed), but reasons for their invalidity are not given. Feedback is given at the end of the problem, letting the user know if the solution was fully simplified or if the minimum number of steps was taken, but no guidance is given on how to improve if those are not the case.

Provides scaffolding: 3

The complexity gradually increases, and new components are introduced one at a time. During the solving of a problem, however, no assistance is provided to prevent a user from getting stuck.

Provides opportunities for reflection: 2

No history of actions. No sense-making questions. The user is informed when the problem could have been solved more efficiently or further simplified, which could conceivably be an opportunity for reflection, but is treated more like a score, and no real motivation is provided for the student to try and improve this. An undo button is present, but there is no indication that it should be used for review rather than to correct mistakes.
Easy to use: 5

Even with no prior algebraic skills, a user would likely be solving problems a very short time after starting the app. Instructions are kept to a minimum, and the app is designed for the users to immediately start completing levels and learn as they go. Initial levels are very basic, and concepts are introduced one at a time, through play, rather than lengthy help files or text.

Documentation is clear and easy to understand: 3

Documentation is limited to in-problem text and icons the first time a concept is introduced. This text is clear, and is linked to the problem currently being attempted. However, once that problem is finished, there is no way of determining in which problem a concept was introduced, so it is difficult to get help on a particular concept if one needs a review. Outside of this in-problem text, there is no documentation whatsoever.

Interface is clear, consistent, and follows technology standards: 5

Object placement is consistent from problem to problem. All buttons are clearly identified. Text is easily readable. Standard icons are used for actions such as undo and restart.

Usable by a variety of learners: 5

Students can solve problems with no prior algebra knowledge. Problems vary in complexity. Advanced users can attempt to get three stars on each problem. Multiple languages are available.
Numerosity: Play with Math!, Chapter 5

Numerosity is the first app developed by ThoughtBox, a company dedicated to “gameful learning”. The development of ThoughtBox and Numerosity is partially supported by a grant from Ireland's National Digital Research Centre. Numerosity is targeted at children aged 8 to 11.

Numerosity is an app that presents the user with an expression or an algebraic equation, and four possible number choices. The user must drag a choice onto an operation that results in that number, or drag a choice onto an unknown that makes the equation true. If the user drags a choice onto the correct operation, the operation and the two operands are replaced with the choice. If the expression is not fully simplified or the equation not solved, the user is presented with four new choices. The user can also drag terms around to rearrange them, which also results in the user being presented with new choices. This repeats until the expression is fully simplified or the equation is solved. A typical problem is shown in Figure 11.

Numerosity: Play with Math! contains 5 chapters, each concerning a different topic in math. Chapter 5 is entitled “Negatives and the Order of Operations”. Despite this title, the last three lessons concern solving for an unknown in an equation. As this chapter is the only one involving algebraic concepts, the review will be limited to that chapter.

In March 2013, Numerosity: Play with Math! was removed from the app store by the developer, who indicated that five new apps will replace it, one for each chapter. As of April 20, 2013, only Numerosity: Play with Addition! (a replacement for Chapter 1) was available.
Meaningfully interactive: 5

The user is continually in control of what is happening on-screen. The user transforms the displayed equation through his or her actions, by rearranging terms and substituting numbers in place of operations and operands.

Natural actions on objects: 3

Terms of the equations can rearranged by the congruent action of dragging. It is not always immediately clear, however, the boundaries of the term that is being interacted with.

Figure 11. The user must drag the 5 onto the division operator or the 16 onto the unknown.
In addition, simplifying expressions requires choosing from a list of numbers outside the expression, rather than actions on the expressions themselves.

Visible consequences of actions: 2

When dragging terms to the opposite side of the equals sign, the operator changes without any additional indication. Terms can be dragged anywhere on screen, and they will “jump” without any animation to the nearest location in the expression without any indication as to where they are going. The rest of the equation shifts immediately to make room for the term and to close the gap where the term originated. When dragging terms, the terms change size, making them seemingly not fit the gap that they came from. When multiple terms are simultaneously dragged to the opposite side of the equation, all the operators change simultaneously and instantaneously, making it difficult to discern which operators changes and which did not.

Range of actions not limited by technology: 2

If 8*x is on the left side of the equation, the 8 can be dragged to other side of the equation, but not if 8*x is on the right side. The user can not perform the same operation to both sides, only drag terms around and substitute numbers in to simplify. Only four possible numbers are presented as choices to substitute; not all valid choices are always presented.

Necessitate thoughtful, deliberate actions: 2

The final solution is always one of the four possible choices; after the first step, the user has seen two lists of choices, and the correct solution is most often the only choice that has been repeated. In addition, the final solution is almost always one of the two choices that are smallest in magnitude. Choosing the smallest choice in magnitude immediately after the
problem is presented will result in a majority of correct answers. In addition, there is very little penalty for incorrect answers, so a student can very easily and very quickly guess his or her way through a large number of problems.

**Cognitively faithful interactions:** 2

A student is unlikely to understand how operators change when dragging terms since the change is not clearly visible. The app treats subtraction and negative numbers interchangeably at times, which is likely to lead to confusion (e.g., the user is expected to drag -9 onto the second subtraction operator in the expression 8 - 5 - 4 to simplify it to 8 - 9).

**Mathematically faithful:** 1

In an equation such as \(-3 \ast x = -9\), the 3 can be dragged to the other side, creating the “equivalent” equation \(-x = -9 \ast 3\). The equation can then be simplified to \(-x = -27\), but the “correct” choice for x remains 3. In the equation \(x = (17 + 9) \div 2\), the entire term on the right side of the equality can be dragged to the left side to create the supposedly equivalent equation \((17 + 9) \div 2 + x = 0\). From here, the equation can be simplified and rearranged to \(x = -13\), and yet, the app rejects -13 as a correct solution. Actions are not always reversible when they should be. Terms that should be separable are not always. The app can only display integers, but problems can involve division that does not result in integers, in which case either no valid choices are presented or the app accepts the integer that is closest to the quotient, as in Figure 12. Equations can often be transformed into things which are no longer sensible equations, containing mismatched parentheses or operands without operators, as shown in Figure 13.
Figure 12. No valid solution is provided for the division.

Figure 13. An equation that no longer makes sense
Cognitively Demanding: 2

Solving a problem generally involves repeating two steps, in some order: dragging terms, and performing arithmetic. Neither one requires much critical thinking in this app.

High-quality presentation: 2

Visible and audible effects indicate correct and incorrect answers. Music can be turned off if distracting. Parts of the equations are sometimes slightly mis-aligned. The equals sign and negative signs can be tapped on, causing them to spin, an action that has no mathematical meaning. Color are used for the different operations, but adds little pedagogical value. When a multiplication table is requested by the user, it often covers up the equation.

Provide relevant and timely feedback: 2

The only noticeable feedback is whether the user's choice was correct or incorrect. The app does not explain why the choice was wrong (e.g., the order of operations was violated).

Provides scaffolding: 3

Lessons gradually increase in complexity. A hint button is provided; when clicked, it highlights an operation that can be performed next, and shows a portion of a number line or a multiplication table, depending on the operation highlighted. The multiplication table, however, only shows multiplicands up to 12, even if the operation is being performed on higher multiplicands. The number line may be partitioned into segments of length other than one, which can make it difficult to use for addition or subtraction.
Provides opportunity for reflection: 1

No history of actions. Users are encouraged to answer as quickly as possible to move on to the next problem and obtain a higher score. No questions or tasks to tie concepts together.

Easy to use: 5

Any student familiar with the rules for the order of operations should be able to begin using the app without much instruction or technical knowledge.

Documentation is clear and easy to understand: 2

The entirety of the documentation consists of a single animation lasting approximately 5 seconds.

Interface is clear, consistent, and follows technology standards: 4

Buttons, numbers, and operators are clearly visible. The equals sign often changes location during the solving of a problem. Buttons and menus act as expected.

Usable by a variety of learners: 5

The app starts out with very simple addition problems, or users can choose to start at higher levels. Advanced users can aim for high scores and rewards. Multiple approaches can sometimes be taken to solve a problem.

touchyMath

touchyMath is the first app developed by Joel Martinez, intended for visual learners making the jump from arithmetic to algebra.
touchyMath allows the user to solve randomly-generated linear and quadratic equations of a single variable, and systems of two linear equations with two variables, or enter his or her own equation to attempt to solve it. One of these randomly-generated problems is shown in Figure 14.

![Figure 14](image)

*Figure 14. A randomly generated problem*

**Meaningfully interactive: 5**

The user is in control of the equation on-screen. He or she can transform it through a variety of actions and gestures. Every lesson is accompanied by interactive examples that the app walks the user through. An example lesson is shown in Figure 15.
Natural actions on objects: 2

Negating both sides is accomplished by double-tapping on the background. Non-essential parentheses are removed by pressing and holding on the term inside the parentheses. Dividing both sides of an equation requires a two-finger tap on an individual term. All of these physical actions are either away from the object that the action is performed on, or incongruent to how a user would perceive the mathematical action being performed.

Figure 15. An interactive lesson on solving linear equations
Visible consequences of actions: 5

When negating both sides of an equation, terms are highlighted and negated one term at a time. When substituting one equation into another, the substituted parts of the latter equation are colored differently. When dragging terms, gaps are made indicating where the term will be placed upon release. Audible tones accompany visual changes to indicate when terms will become negative or positive while dragging. When an equation is multiplied or divided by a number or expression, each term in the equation highlighted and divided one at a time.

Range of actions not limited by technology: 4

The actions available are context-sensitive (e.g., activating the term $\frac{2x}{7}$ will only let the user divide the equation by 2 or x, or multiply the equation by 7). Any actions that would reasonably help solve the equations, however, are generally available.

Necessitate thoughtful, deliberate actions: 5

Since the user must select both the scope of the action and the action itself, random use is unlikely to result in finding the solution.

Cognitively faithful interactions: 2

$x = 7$ is considered a solved equation, but $7 = x$ is not. Multiplying 3 by 5 results in 15, but multiplying 4 by 4 results in 4², which can then not be added to another integer until the user simplifies 4² into 16. Adding $-1 \ast 7$ and $1 \ast 4$ results in $-3 \ast 1$, that while correct, is likely to be confusing, particularly because the order of the operands is switched. Adding...
\[-\frac{2x}{7} + \frac{2}{7}\] results in \[\frac{2(-x+1)}{7}\], which could likely be confusing for some students. Some non-essential parentheses are not treated as such (i.e., they can not be removed like other non-essential parentheses can be). In some of the examples, variables are treated interchangeably with units. The app states that “adding 5 chocolates to 25 chocolates will result in 30 chocolates”, while displaying the expression \[5 \ast \text{chocolates} + 25 \ast \text{chocolates},\] which reinforces a common misconception that many students have (Booth, 1988). One of the equations generated by the app had no solutions, but the user is still expected to try and solve it, even after transforming it to an equation such as \[2 = 0.\] In an equation, the equals sign can be dragged and placed between any terms. While the resulting equation is accurate, this could certainly obscure the properties of equality.

Mathematically faithful: 4

The computer algebra system is to handle very complex expressions with accuracy. All roots are considered to be positive, so unless the user is using the build-in quadratic equation solver, only one solution can ever be found.

Cognitively demanding: 4

The later problems and lessons have multiple steps and incorporate multiple topics taught previously in the app. Solving problems simply involves performing procedures that have been learned, but choosing the appropriate procedure possesses some cognitive demand.

High-quality presentation: 5

Colors are used throughout to highlight various aspects of the equations. Parts of the equation fade out when the user is working on a particular term, allowing the user to focus on
the term being worked on. Animations are used to direct the user's attention as mathematical actions are being performed. When a term is selected, it is highlighted so the user knows the boundaries of the term. Sound effects are used to indicate changes as well as indicating invalid moves.

*Provides relevant and timely feedback: 5*

When invalid actions are attempted, the app highlights terms or operations that indicate why the action was invalid. When a user submits an unsimplified solution, the app will indicate what parts of the solution can still be simplified.

*Provides scaffolding: 4*

Concepts are introduced one at a time. The later topics in particular build on previous topics. Examples are given to the user to work through for each concept, with instructions and mathematical explanations for each step. Outside of the examples, however, there is no option to choose the complexity of the problem, other than choosing the type of equation (e.g., linear, quadratic, system of linear), and hints for these problems are not given.

*Provides opportunity for reflection: 4*

An extensive history of actions is provided, as shown in Figure 16. Some examples are given of real-life applications in which strategies learned can be applied, along with a discussion of the mathematics involved. No questions are presented to the user, however.
Easy to use: 2

Other than dragging terms, most actions are not intuitive. In particular, having to double-tap within and outside of terms to select the scope is not something that is immediately clear. It would likely take several minutes of experimentation to figure out how to solve a problem. Most users would likely have to consult the documentation for at least some of the actions.

*Figure 16. The user’s history of actions after solving a linear equation*
Documentation is clear and easy to understand: 5

Documentation is clear and well-organized, with interactive examples. Video tutorials are also given for the more common actions.

Interface is clear, consistent, and follows technology standards: 3

A number is ever-present in the background – presumably a score of some type, though this is never explained or used. Menus are easy to navigate and buttons are clearly identified and act as expected. Equations can often exceed the boundaries of the screen, and while the user can pan, there is no zoom functionality that would allow the entire equation to be viewed at once.

Usable by a variety of learners: 3

No varying difficulty levels or complexities. Multiple solution strategies are only sometimes possible.

Cover Up

Cover Up is one of several apps developed by the Center for Algebraic Thinking, a consortium of four Oregon universities. The Center for Algebraic Thinking is funded by the U.S. Department of Education Fund for the Improvement of Post-Secondary Education.

Cover Up lets users solve algebraic equations by operational unwinding. Starting with an equation such as $3(x^2+1)-4=47$, the app “covers up” part of the equation, resulting in the one-step equation $\text{■} - 4 = 47$, as seen in Figure 17. Using a calculator-type interface, the user enters the number that makes the one-step equation true, then a new equation is created with the expression that is covered up equal to the number the user entered,
\[ 3(x^2 + 1) = 51 \] Part of this equation is covered up, \( 3 \square = 51 \), and the process repeats until the original equation is solved.

![Equation before and after it has been "covered up".](image)

**Figure 17. An equation before and after it has been "covered up"**

**Meaningfully interactive:** 3

Given a one-step equation, the user enters the answer that will create a new equation, but the deciding which part to cover up and creating a new equation is done automatically by the app. The user never interacts directly with the equation, and the user can not determine how the equation will be transformed.

**Natural actions on objects:** 2

All actions performed by the user are on a calculator-type interface. Any actions performed directly on the equations are done by the app. The app implies a physical motion...
(covering up part of the equation) that has an analogous mathematical action, but this motion is never actually performed by the user.

Visible consequences of actions: 3

The part of the equation that will get covered up next is clearly identified. However, it is not clearly shown how the user's input is used to create a new equation.

Range of actions not limited by technology: 2

The user is limited to choosing the solution to the current one-step equation that is presented. Users can not choose what part of the equation to cover up. In addition, there is a very small selection of problems in the app, so the user will be repeating the same procedures after only a few problems.

Necessitate thoughtful, deliberate actions: 3

The user has to solve each one-step equation before he or she is able to proceed. Many actions that could require deliberation are performed automatically by the app. All solutions are integers and usually small in magnitude, and there are no penalties for incorrect answers, so it is conceivable that a student could simply guess until they found the correct solution.

Cognitively faithful interactions: 3

After using the app, the user should better understand how a complex equation can be decomposed into smaller one-step equations. However, the focus is on how to perform the procedure, rather than why it works. There is no discussion of what types of equations this procedure will work on, or on how to choose what parts of the equation to cover up.
Mathematically faithful: 3

For equations with more than one solution, both solutions are only sometimes accepted. For example, if the equation displayed is \[ (■)^2 = 4 \], the app will not accept -2 as an answer. Even when both solutions are accepted, the app will often not accept answer such as \(3 ± 4\), even if the equation displayed is \(■ − 3 = ±4\). However, if the user enters either 7 or -1, the app will accept it, and then display the correct answer as \(3 ± 4\). If \(±4\) is an accepted answer, then \(0 ± 4\) will not be.

Cognitively demanding: 2

The only thing the user has to do is solve one-step equations such as \(■ + 2 = 3\). All the other work, including deducing which equations to solve, is done automatically by the app.

High-quality presentation: 2

Any colors and animations used do not benefit learning. There are no sounds or videos, and very little graphics to support learning.

Provides relevant and timely feedback: 2

The only feedback is “Correct!” or “Try Again”

Provides scaffolding: 3

The user can choose the level of complexity of the equation, gradually increasing difficulty if he or she wishes. Within each problem, little scaffolding is provided, although many of the intermediate steps are done automatically, and the typical student would not be in danger of getting stuck.
Provides opportunities for reflection: 3

The user can back up a step at any point. After a problem is solved, the app displays a series of steps showing answer being substituted back into the original equation and the two sides of the equation evaluating to the same number. There are no sense-making questions asked of the user, and no links are made to larger algebraic concepts.

Easy to use: 5

At each step, it is very clear what the user is to do and how to do it. No special technical skills are needed, and users should be able to complete problems without any instruction.

Documentation is clear and easy to understand: 3

The documentation consists of a single paragraph of text. While the paragraph itself is clear, no pictures or examples are given. The help text can not be accessed from within a problem.

Interface is clear, consistent, and follows technology standards: 3

Most buttons are clearly identified and act as expected. Some buttons are extraneous, such as the space key and the fraction bar, which are never used in a correct solution, since all answers are integers. The calculator interface can vary in size, depending on the equation presented. At each step, the equation is displayed twice, when once would suffice.

Usable by a variety of learners: 3

The user can choose the complexity of the problem. Once starting a problem, however, the user is forced into a single strategy and a single approach.
Mathination - Equation Solver

Mathination - Equation Solver is the sole app from developer Orion Math. Mathination lets the user solve algebraic equations through touch and gestures. The user can rearrange terms, factor and expand polynomials, perform arithmetic operations, and transform the equation by performing actions such as negating both sides or taking the reciprocal of both sides. The app includes a series of lessons intended to teach the user the range of gestures, and includes several practice problems for each lesson. One such lesson is shown in Figure 18. Users can also enter their own equation and attempt to solve it.

![Factoring Polynomials — Factoring $x^2 + bx + c$](image)

Factor and solve for $x$.

$x^2 - 7x + 10 = 0$

Tip: Squeeze all three terms of the polynomial to factor it. After factoring the polynomial, apply the "zero product" rule to solve the equation.

Figure 18. A lesson on solving quadratic equations by factoring
Meaningfully interactive: 5

The user is continually in control of the equation shown on-screen. The user can mathematically transform the equation by squeezing, expanding, and dragging terms.

Natural actions on objects: 3

Most actions are performed with gestures on the objects being performed on. Several transformations, though, require picking an option from a menu. Some gestures are incongruent with their mathematical action, and different actions require the same or similar gestures. For example, squeezing terms together is the gesture both for adding and for factoring. Multiplying single terms requires that the user make a squeezing gesture, but multiplying polynomials requires the reverse (an expanding) gesture.

Visible consequences of actions: 2

Changes in equations happen instantly, even if user is still continuing the gesture.

\[
\frac{(7+2)}{3}\]

Multiple actions often happen simultaneously (e.g., \(\frac{30}{s(s-2)} = -2\) is replaced by 3 when performing the gesture to add 7 and 2). Distribution, factoring, multiplying polynomials, and adding fractions with different denominators are all multi-step processes in which all the steps happen simultaneously. Changes are not highlighted in any way.

Range of actions not limited by technology: 2

In the equation \(\frac{-30}{s(s-2)} = -2\), the s can be dragged to the other side, but not the \(s-2\). In the equation \(-3x = -9\), the user cannot divide by -3 – he or she must first divide by 3, and then negate both sides using a menu option. Cube or higher order roots cannot be
taken; \( x^3 = 8 \) can not be solved. Likewise, the user can square both sides, but not other powers. Two solutions can be found for \( x^2 = 4 \), but in \((x^2 - 4)(x^2 - 9) = 0\), the solutions can not be simplified further than \( x^2 = 4 \) or \( x^2 = 9 \). The user can perform only a small range of operations (negation, squaring, square root, and reciprocal) to both sides of the equation. In custom problems, decimals can not be entered.

*Necessitate thoughtful, deliberate actions: 3*

Most problems, particularly the sample problems, are easily solvable by random use of the three basic gestures (squeezing, expanding, and dragging) and the use of the undo button.

*Cognitively faithful interactions: 2*

In the sample problems, only the exact answer expected will be accepted. For example, one problem will accept \( \frac{(a+b)}{2} \), but not \( \frac{(b+a)}{2} \). Terms must be in order of decreasing degree to be accepted. For custom problems, the app can not recognize when the problem is solved. Hints can be provided when attempting to factor a quadratic, but are sometimes confusing and even inaccurate, as shown in Figure 19. It is often not clear what actions have just been performed, or how the equation has changed. In the lessons, the only thing that is taught is how to perform gestures and execute a procedure – no mention is made of the math behind the procedure or why the procedure is being performed.
Mathematically faithful: 4

All actions that are able to be performed seem to be valid, along with their resulting equation. When providing hints when factoring quadratics, the factors of $4x^2$, for example, are given as $x$, $2x$, and $4x$ (i.e., only first-degree factors are given).

Cognitively demanding: 3

In the sample problems, each problem is slightly, but significantly, different than the last, requiring different procedures. However, each sample problem outlines the procedure to take, and except for some of the problems near the end, solving each problem simply requires
following directions and applying the stated procedure. Custom problems could potentially have a higher cognitive demand.

*High-quality presentation: 3*

In the factoring pop-up, the buttons are not always visible. The history screen is jumbled, with overlapping text, if the problem has been split into two or more equations, as shown in Figure 20.

![Figure 20. The history is jumbled and difficult to comprehend](image)

*Provides relevant and timely feedback: 2*

When choosing terms to factor a quadratic, the product of the terms is displayed for each choice. Audible tones indicate invalid actions, but only sometimes, and reasons the
action is invalid are not given. Outside of this, the only noticeable feedback is telling the user that a sample problem has been completed. In custom problems, no feedback is given, either during the problem, or on completion.

*Provides scaffolding:* 4

The sample problems increase in complexity and build on each other. A hint button is provided when factoring a quadratic, listing factors, and suggesting how to select factors from these pairs.

*Provides opportunities for reflection:* 3

A history of actions is provided. An undo button is provided, but is seemingly meant only for correcting mistakes. Links are not made between problems, and sample problems end as soon as a solution is found.

*Easy to use:* 3

The app often has trouble differentiating between and recognizing gestures. Simple gestures can often result in significantly more complex equations. Some actions are difficult to perform or are hidden in menus.

*Documentation is clear and easy to understand:* 4

Most documentation is within the sample problems, but are fairly easily located and accessed. Opening a sample problem, however, will halt the problem the student is working on, losing all progress.
**Interface is clear, consistent, and follows technology standards: 3**

No pan or zoom, even though problems can run off-screen. Location of elements is consistent, and entering custom mathematical problems and formulas is easy with the built-in keyboard.

**Usable by a variety of learners: 3**

Most of the earlier sample problems have only one viable approach. Later problems and custom problems, however, potentially have several solution strategies. Although sample problems gradually increase in complexity, these are not necessarily more cognitively demanding, and users are expected to all start at the first problem.

**Cross-app Analysis**

By examining scores across apps, we can observe commonalities among apps, and determine what this group of apps as a whole does well, or does poorly. The highest average score by far (4.57) is in the criteria *meaningfully interactive*. One of the steps in the process of selecting the apps to review, however, was filtering for a certain level of interactivity, so there is a selection bias influencing the scores in this criteria. After *meaningfully interactive*, the four criteria with the highest average scores are *usable by a variety of learners* (4.14), *interface is clear, clean, and follows technology standards* (4.00), *visual consequences of actions* (3.86), and *natural to use* (3.86). Three out of these four are three of the four criteria that make up the *Usability* category.

The lowest average scores, on the other hand, are for the criteria *provide opportunity for reflection* (2.43), *range of actions not limited by technology* (2.57), *cognitively faithful interactions* (2.71), and *provide relevant and timely feedback* (3.00). These are areas in
which developers may want to devote extra attention to when developing future apps of this type.

_Correlation with App Store ratings_

For each of the 16 criteria, the Pearson product-moment correlation coefficient was calculated for the App Store ratings against the scores in that criteria for each of the seven apps. The correlation coefficients can be seen in Table 4. At the p < .02 level, the scores for three criteria are significantly positively correlated with the App Store ratings: _natural actions on objects_ \((r = 0.895, p < 0.01)\); _interface is clean, clear, and follows technology standards_ \((r = 0.895, p < 0.01)\); and _usable by a variety of learners_ \((r = 0.850, p < 0.02)\). The average score across the _usability_ category was also significantly correlated with the App Store ratings. This would indicate that, of the sixteen criteria in this instrument, these three criteria (and the _usability_ category in general) may be the most valued for a typical user.

Several criteria, however, particularly in the _quality of content_ and _feedback and support_ categories, have near-zero or moderately negative correlations with App Store ratings. This suggests that typical users have significantly varying criteria for evaluating apps, or that they are evaluating these apps based on criteria not found in this instrument. It is unknown whether these users are educators or students (or perhaps neither), and as discussed in Chapter 2, research (Akpinar, 2008; Kay & Knaack, 2008b) suggests that when evaluating educational tools, teachers and students tend to place higher values on different areas – some of which are unrelated to the learning benefit the app provides. Indeed, a cursory examination of some of the reviews in the App Store suggests that some of these users may be evaluating these apps on criteria other than those that contribute to the learning of the mathematical
Table 4. Pearson correlations with App Store ratings.

<table>
<thead>
<tr>
<th>Criteria/Category</th>
<th>Correlation with App Store Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meaningfully Interactive</td>
<td>.072</td>
</tr>
<tr>
<td>Natural actions on objects</td>
<td>.895*</td>
</tr>
<tr>
<td>Visible consequences of actions</td>
<td>.262</td>
</tr>
<tr>
<td>Range of actions not limited by technology</td>
<td>-.433</td>
</tr>
<tr>
<td>Necessitate thoughtful, deliberate actions</td>
<td>-.138</td>
</tr>
<tr>
<td>Cognitively faithful interactions</td>
<td>.620</td>
</tr>
<tr>
<td><strong>Interactions category overall score</strong></td>
<td>.399</td>
</tr>
<tr>
<td>Mathematically faithful</td>
<td>.033</td>
</tr>
<tr>
<td>Cognitively demanding</td>
<td>.016</td>
</tr>
<tr>
<td>High-quality presentation</td>
<td>.196</td>
</tr>
<tr>
<td><strong>Quality of content category overall score</strong></td>
<td>.105</td>
</tr>
<tr>
<td>Provides relevant and timely feedback</td>
<td>-.211</td>
</tr>
<tr>
<td>Provides scaffolding</td>
<td>-.478</td>
</tr>
<tr>
<td>Provides opportunities for reflection</td>
<td>-.771</td>
</tr>
<tr>
<td><strong>Feedback and support category overall score</strong></td>
<td>-.541</td>
</tr>
<tr>
<td>Easy to use</td>
<td>.561</td>
</tr>
<tr>
<td>Documentation is clear and easy to understand</td>
<td>-.388</td>
</tr>
<tr>
<td>Interface is clear, consistent, and follows technology standards</td>
<td>.895*</td>
</tr>
<tr>
<td>Usable by a variety of learners</td>
<td>.850*</td>
</tr>
<tr>
<td><strong>Usability category overall score</strong></td>
<td>.863*</td>
</tr>
<tr>
<td>Total overall app score</td>
<td>.191</td>
</tr>
</tbody>
</table>

*Significant at the p < 0.02 level
content (e.g., price, ability to run on their specific device, presence of in-app purchases). A further study could be conducted that examined the App Store reviews (rather than just the numerical ratings) and coded them based on the criteria in this evaluation instrument, perhaps revealing criteria of a different nature than those included in this instrument.

Reliability of the instrument

To assess the reliability of the instrument, a graduate student in mathematics education was asked to rate two of the apps – *Algebra Touch* and *Mathination - Equation Solver* – using the process outlined in Chapter 3. The two-way, mixed intraclass correlation coefficient (ICC) for absolute agreement across all 16 criteria was calculated, and can be seen in Table 5. Supposing that the (arbitrary) threshold for sufficient reliability is 0.75, the combined ICC value of 0.69 falls slightly under this threshold. Possible methods to increase reliability are discussed in Chapter 5. More evaluations with additional raters would have to be performed to determine the reliability of each individual criteria.

*Table 5. Intraclass correlation coefficient*

<table>
<thead>
<tr>
<th>App</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algebra Touch</td>
<td>.790</td>
</tr>
<tr>
<td>Mathination</td>
<td>.638</td>
</tr>
<tr>
<td><strong>Combined</strong></td>
<td><strong>.686</strong></td>
</tr>
</tbody>
</table>
CHAPTER 5
DISCUSSION AND CONCLUSIONS

In the process of reviewing these apps, several details regarding the instrument and the apps surfaced that warrant further discussion.

The criteria

In choosing the apps to review, one of the steps was filtering for interactivity. If the app was non-interactive, then most of the criteria (e.g., visible consequences of actions, cognitively faithful interactions) would be inapplicable. For that reason, meaningfully interactive should perhaps be a requirement to employ this evaluation instrument, rather than a criteria within it.

One of the lowest-scoring criteria across the seven apps is range of actions not limited by technology. It is interesting to note that some of the apps examined (e.g., DragonBox+ Algebra, Cover Up Math) intentionally limit the mathematical actions available to the user, presumably to benefit the user's learning. Further research would have to be conducted to determine if approaches like this are effective, and if so, this criteria would likely need modification.

When considering the documentation is clear and easy to understand criteria, no explicit differentiation was made between documentation pertaining to mathematical concepts and documentation pertaining to the technical aspects of the app. In many of the apps, the two are intertwined, as the mathematical concepts are taught alongside the gestures needed to perform them. At least one of the apps, however (DragonBox+ Algebra), included almost no documentation of the mathematical type, presumably as a way to let users
construct their own mathematical knowledge as they work through the app. As this approach is widely advocated, it could be argued that this criteria should be modified only to evaluate the documentation as it pertains to how to use the app. If limited mathematical documentation for the user is appropriate, then there may be a need for a separate documentation aimed at the teacher, one that discusses the mathematical concepts within and how the app helps the students learn these concepts.

One of the difficulties in applying the *natural actions on objects* criterion is deciding if a physical action is congruent with a mathematical concept without knowing what the typical user's conception of the mathematical concept is. For example, when multiplying two polynomial expressions, is dragging the polynomials together to “combine” them more congruent than dragging them apart to “expand” the expression? Is tapping on an addition sign to perform the addition less congruent than dragging the two addends together? Does this change if the operation is subtraction? A topic for further research could be determining which actions students feel are congruent with their perceptions of mathematical concepts.

One of the things not addressed in this instrument is the stability of the app. One of the apps in particular, Numerosity, crashed several times during the review. Despite not being directly related to algebra, the 27 experts that gave feedback for Bokhove's and Drijver's (2010) instrument for evaluating digital algebra tools, on average, gave the highest priority to the stability and performance of the tool. It is reasonable to believe that regardless of the pedagogical quality of an app, having it frequently crash could seriously hinder its value as an effective learning tool.
The apps

We can split the seven apps into two groups, based on overall score from the instrument. The four “high” apps – Algebra Touch, touchyMath, Dragonbox+ Algebra, and Hands-on Equations, with scores ranging from 3.63 to 4.15 – and the three “low” apps – Mathination, Cover Up, and Numerosity, with scores ranging from 2.58 to 3.10.

The main source of differentiation between the two groups came from two categories: interactions and quality of content. The apps in the high group scored very well in these two categories, while apps in the low group scored poorly. These also happen to be the two categories in which an app that scored well in would be an app that would most likely help students overcome the common difficulties in algebra – by helping them understand the parts of an equation, and by promoting formal methods and precise notation.

Most of the apps reviewed in the high group were significantly lacking in a single area. Hands-on Equations and Dragonbox+ Algebra, for example, scored a 3.8 or higher in every category except for feedback and support, where they scored a 2.3 and a 2.7, respectively. touchyMath, on the other hand, received the highest rating among all the apps in the feedback and support category, but was tied for the low score with a 3.3 in the usability category. What this means, however, is that with improvements in just that one area, these apps could perhaps become ones that have great potential in the classroom.

Instrument Reliability

The intraclass correlation coefficient of 0.69 for the reliability of the instrument fell slightly under what might arbitrarily be considered an acceptable value of 0.75. Several possible steps could be taken to increase reliability. First, the second rater received no
training or participated in any discussion regarding the instrument before evaluating the apps. It is reasonable to assume that receiving training, participating in discussions regarding the criteria, or even simply reviewing a previous evaluation and justifications for scores (e.g., an evaluation as presented in Chapter 4) would increase reliability. Second, as discussed in Chapter 2, Vargo et al. (2003) showed that collaboration with other raters during an evaluation can significantly increase reliability of the evaluation instrument being used. Third, it could be required that multiple raters be required to evaluate an app, and the average score of the evaluations be used. Fourth, some criteria in the instrument may be vague or unclear, and clarifying or modifying them could increase reliability. To determine which criteria, however, more evaluations with additional raters would have to be performed to determine the reliability of each individual criteria.

Connections to Other Research Involving These Apps

Ottmar, Landy, and Goldstone (2012) used a modified version of the Algebra Touch app as part of an intervention encouraging students to treat symbol systems as physical objects that move and change. Seventy eighth-grade students participated in a study in which they were given a whole-group lesson on simplifying algebraic expressions and were then asked to simplify expressions first using colored tiles, and then using the Algebra Touch app. The amount of problems solved without structural errors significantly increased (from 9.4% to 54%, p < 0.01) after this intervention. As Algebra Touch received the highest overall score among the seven apps evaluated in Chapter 4, this could suggest that the instrument is a good predictor of the learning benefit of an app. Clearly, though, more research would have to be done with other apps to substantiate that claim. Interestingly, neither mathematical anxiety or...
pre-test performance were predictors of post-test performance in Ottmar, Landy, and Goldstone's study, perhaps speaking to the usable by a variety of learners criteria, for which Algebra Touch received the maximum score of 5 in the evaluation reported in Chapter 4.

Although no studies were located involving the Hands-on Equations iPad app, there have been studies involving the physical manipulatives that the app emulates. Suh and Moyer (2007), for example, examined the effects that using the physical Hands-on Equation manipulatives had on learning outcomes of students. They also examined a virtual balance learning object that very closely resembles the Hands-on Equations iPad app, but provides more frequent feedback and support, an area in which the Hands-on Equation app was lacking, according to the evaluation reported in Chapter 4. During a week-long unit focusing on algebraic relationships, one group of students used the physical manipulatives, while the other group of students used the digital learning object. Both groups saw significant (p < 0.001) gains in scores on the post-test from the pre-test. As the Hands-on Equations iPad app closely emulates the physical manipulatives, and bears many resemblances to the learning object used in this study, it is a reasonable assumption that the Hands-on Equations iPad app would also benefit learning.

Additionally, Borenson and Barber (2007; 2008) have conducted several studies examining pre-test, post-test, and retention test scores of fourth-, fifth-, sixth-, and eighth-grade students after using the physical version of Hands-on Equations. In nearly all of their studies, they saw significant gains in scores from the pre-test to the post-test, as well as from the pre-test to the retention test.
None of the studies described in this section contained a control group or compared
gains to those of students in more traditional learning environments, so while there is
evidence that Algebra Touch and Hands-on Equations (at least, the physical version) benefit
learning, whether they are satisfactory alternatives to other options used in the classroom has
not been fully answered. However, since algebra is one of the most troublesome classes for
many students, anything that significantly benefits algebraic learning has the potential to be a
useful educational tool.

No other studies were located involving any other of the seven apps evaluated in
Chapter 4. To fully assess the validity of this instrument and determine if a significant
correlation exists between ratings and learning outcomes, more research would first have to
be performed to determine the effect that the use of these apps (or other apps evaluated with
the instrument) has on learning outcomes.

Limitations of the Instrument and Other Areas of Consideration

The purpose of this instrument is to aid in the selection of appropriate iPad apps that
can potentially aid in the learning of mathematics. However, the selection of an app is only
the first step. How these apps are used can have as large an impact on a student's learning as
which app is used (Laborde, 2007; Polly, 2011). Another entire instrument could be created
on how to effectively use these apps. Teachers must pose mathematically rich tasks
(Laborde, 2007), focus on the math rather than the technology, support collaborative
environments in which students can communicate their ideas (Polly, 2011), intervene at
critical times (Hoyle & Noss, 2003) and so on.
Even if educators are provided with tools to easily find and select iPad applications for their mathematics classroom, there are a number of other factors that could prohibit the widespread use of these apps. Teachers are often unfamiliar with iPads and generally need training on how to effectively use them in the classroom (Crichton et al., 2012). Even with training, however, integrating new technology into the classroom is not an easy task for teachers. Instead of following a textbook, they must design tasks and worksheets, manage several different kinds of time in their classrooms, and find ways to relate paper-and-pencil techniques to technology-assisted ones (Polly, 2011). Goos and Bennison (2002) found that even when teachers work in technology-rich environments, have had ample training, and are encouraged to use technology by the administration, they often do not develop a proclivity to use the technology.

So, while there are many factors that can ultimately determine the benefit an iPad app will have on students' learning of algebraic equations, this instrument can assist educators in choosing apps that possess qualities that research suggests will promote students' learning and understanding.
REFERENCES


Cromack, J. (2008) Technology and learning-centered education: Research-based support for how the tablet PC embodies the seven principles of good practice in undergraduate


APPENDICES
Appendix A

MERLOT Criteria

Quality of Content:

• Does the software present valid (correct) concepts, models, and skills?
• Does the software present educationally significant concepts, models, and skills for the discipline?

Potential Effectiveness as a Teaching-Learning Tool:

• Does the interactive/media-rich presentation of material improve faculty and students' abilities to teach and learn the materials?
• Can the use of the software be readily integrated into current curriculum and pedagogy within the discipline?
• Can the software be used in a variety of ways to achieve teaching and learning goals?
• Are the teaching-learning goals easy to identify?
• Can good learning assignments for using the software application be written easily?

Ease of Use:

• Are the labels, buttons, menus, text, and general layout of the computer interface consistent and visually distinct?
• Does the user get trapped in the material?
• Can the user get lost easily in the material?
• Does the module provide feedback about the system status and the user's responses?
• Does the module provide appropriate flexibility in its use?
• Does the learning material require a lot of documentation, technical support, and/or instruction for most students to successfully use the software?

• Does the material present information in ways that are familiar for students?

• Does the material present information in ways that would be attractive to students?
APPENDIX B

Co-operative Learning Object Exchange (CLOE) Criteria

Quality of Content:

• The content of the learning object is accurate.
• The use of technology is appropriate for this content.
• The content is presented clearly and professionally (spelling/grammar, etc).
• Appropriate academic references are provided.
• Credits to creators are provided.

Effectiveness as a Teaching/Learning Tool

• There are clear learning objectives.
• The learning object meets the stated learning objectives.
• The target learners are clearly identified (academic level addressed/technical ability/demographics).
• There are clear instructions for using the learning object.
• The technology helps learners to engage effectively with the concept/skill/idea.
• The learning object provides an opportunity for learners to obtain feedback within or outside the learning object.
• The author provides evidence that the learning object enhances student learning.1
• Pre-requisite knowledge/skills, if needed, are identified.
• The learning object stands alone and could be used in other learning environments.
Ease of Use/Usability

- The learning object is easy to use (i.e. navigation, user control).
- The author indicates whether the learning object is accessible for learners with diverse needs.
- Technical requirements for the learning object are provided.
APPENDIX C

Learning Object Review Instrument (LORI) Criteria

Version 1.3:

• Presentation: aesthetics
• Presentation: design for learning
• Accuracy of content
• Support for learning goals
• Motivation
• Interaction: usability
• Interaction: feedback and adaptation
• Reusability
• Metadata and interoperability compliance
• Accessibility

Version 1.5:

• Content quality
• Learning goal alignment
• Feedback and adaptation
• Motivation
• Presentation Design
• Interaction usability
• Accessibility
• Reusability
• Standards compliance

*Version 2.0:*

• Content quality
• Learning goal alignment
• Feedback and adaptation
• Motivation
• Presentation Design
• Interaction usability
• Accessibility
• Reusability
• Standards Compliance
APPENDIX D

Le@rning Federation Educational Soundness Specification Criteria

Learner focus:

Learning object design that reflects the relevant learner profiles based on intended users, including:

• age and stages of schooling
• community and cultural affiliations
• languages and dialects they speak, read and write
• socioeconomic status
• existing skills, knowledge and understanding

Learning object design that enables learners to interact with, organise, represent, interpret and manage the process of learning and the information flow through:

• making choices and decisions
• inquiring, investigating and problem solving
• eliciting and receiving relevant, timely and informative feedback
• interpreting, developing and presenting meaningful and useful products
• applying knowledge in a range of contexts.

Learning object design that makes explicit and consolidates the process of learning through:

• structuring informational content in order to scaffold student learning
• enabling students to engage at varying levels of complexity
• reflecting an awareness of the varying educational environments in which learning sequences and objects may be used
• facilitating assessment of how learning is progressing.

Learning object design that contextualises student learning through:
• establishing connections with prior and likely future learning-centered
• making explicit to the learner the connections between the online content and the learning intention
• supporting communication, activity and collaborative action, both online and offline
• allowing the input of data collected offline to directly affect the responses, behaviour and product of the learning object
• using authentic situations where appropriate.

Integrity

Learning object design that:
• accurately represents the ways of knowing and conceptualising the content domain
• uses the language and symbols of the content domain and its ways of representation and supports students in developing and using them
• presents controversial issues with balance and fairness and in accordance with mandated curriculum policies, where these apply
• supports students’ deepening of knowledge within the content domain
• assists the learner with identifying and differentiating between different points of view and perspectives presented
• incorporates content area advice supplied by expert representatives from relevant domains and practice areas.

**Useability**

*Learning object design in which:*

• the purpose, process and intended outcomes of the learning are explicit

• learning and information design is intuitive (that is, the user knows what to do and how to do it)

• the time and effort needed to use it is reasonable

• the medium is exploited to maximise the opportunities for learners to achieve the learning outcomes

• content is constructed in manageable and meaningful concept chunks to facilitate learning.

**Accessibility**

*Online content will be accessible when it:*

• complies with accessibility standards for students with disabilities and for rural and remote communities

• utilises the capacity of multimedia to support student acquisition of standard Australian English or standard New Zealand English

• provides specific language support for students whose first language is not English

• is appealing to and inclusive of students of all genders, socioeconomic groups, ages, races and cultures.
APPENDIX E

Learning Object Evaluation Instrument (LOEI) Criteria

Integrity

• The content of the learning object is accurate and reflects the ways in which knowledge is conceptualized within the domain.

Usability

• Clear instructions for using the learning object are provided.

• The learning object is easy to use (i.e., navigation, user control, visibility of system status).

Learning

• Learning objectives are made explicit to learners and teachers.

• The target learners are clearly identified (academic level/technical ability/demographics) and addressed.

• Pre-requisite knowledge/skills are clear with connections to prior and future learning.

Design

• The technology helps learners to engage effectively with the concept/skill/ideas.

• The learning object structures information content in order to scaffold student learning.

• The learning object provides an opportunity for learners to obtain feedback either within or outside the learning object.
• The learning object stands alone and reflects an awareness of the varying educational environments in which learning sequences and objects may be used by the learner.

Values

• The learning object is appropriate for community and cultural affiliations, including language, dialect, reading and writing.

• Help and documentation files are provided for students and teachers including contextual assistance.

• The design of visual and auditory information enhances learning and mental processes.

• The learning object is accessible to learners with diverse needs.

• The learning object does not require instructor intervention to be used effectively in a mixture of learning environments and learning sequences.
APPENDIX F

Learning Object Evaluation Metric (LOEM) Criteria

*Interactivity*

- Meaningful Interactions
- Overall Control
- Multimedia adds learning value

*Design*

- Consistency (Pages have consistent look and feel)
- Layout (Clear and well organized)
- Labeling (title on menu buttons, words on clickable buttons, any labels used to guide navigation)
- Readability (Look of text)

*Engagement*

- Quality of Feedback (refers to feedback given to user to help him/her progress through the learning object)
- Attractive (Has modern, appealing look)
- Graphics (Not Video)
- Learning Mode
- Motivation
Usability

• Natural to Use (intuitiveness of the interface, easy of use)

• Orientation (Does the user know where he/she is at all times?)

• Navigation Cues (breadcrumb paths, page numbering, coloured buttons to indicate change of state, pop-up boxes or mouseovers)

• Instructions

• Appropriate Language Level (appropriate for the user)
APPENDIX G

Identifying Design Principles in Educational Applets (IDEA) Criteria

Motivation

• Provide a familiar problem context.

• Use second person voice.

• Enable a reward for students early in the problem solving (e.g. provide easy questions or activities they can do successfully)

• Match user expectations for playability for videogame-like activities, interactivity, high-quality graphics, etc.

Presentation

• Question, cover story, and/or introduction should be unwordy, unsuperfluous.

• Proofread text, labels, etc., with target users and age range in mind.

• All other things being equal, use professional conventions for content domain.

• Make the links between representations obvious and warranted.

• Use high-quality graphics and other media (e.g., still graphics, audio, animation).

• Draw attention only to things that support the problem solving.

• Make everything described in the question obvious in the applet; align interactive and noninteractive parts.

Support for problem solving

• Supply a history of actions.

• Allow multiple entry points (e.g., ability, experiences, preferences, styles).
• Support multiple approaches and multiple solution strategies (e.g., questions and/or applet).

• Use dynamic multiple representations appropriately (linked or not-linked, multiple or single sources of control).

• Give students opportunities to make predictions, commit to them, and examine outcomes.

• Reward thoughtful strategic use of the tool more than random use.

• Make a pedagogical decision about whether closure is needed.

• Applet should give appropriate status feedback (say the right thing at the right time in the right way).

• Require a level of accuracy necessary for the problem solving.

• Make effort involved in an activity proportional to the importance of what is needed to solve a problem (aside from programming for accuracy).
APPENDIX H

Designing Interactive Mathematics Criteria

• If possible, model the behaviour of onscreen mathematical objects by observing and analyzing their “natural” mathematical properties; otherwise, invent and standardize behaviours

• make interactive behaviours visible, at all stages of the interaction

• KISS - keep it strictly simple

• Keep the interactivity focused

• Adapt basic CHI principles to mathematical interactions

• Use appropriate onscreen graphical design

• Adapt instructional design principles to onscreen presentations

• Allow the viewer to act directly on onscreen objects

• Don't use animation as a substitute for interaction

• Use “hypertelevision” sparingly

• Allow for multiple use and reuse
APPENDIX I

Digital Tools for Algebra Education Criteria

*Algebra category (contains criteria related to mathematics and algebra in particular)*

- The tool enables the student to apply his or her own paper-and-pencil reasoning steps and strategies
- The tool is mathematically sound and faithful to the underlying mathematical properties (e.g. conventional representations, sound operations)
- The tool’s openness enables the student to express mathematical ideas and strategies
- The tool has the ability to randomize algebra assignments
- The tool has the ability to combine questions into larger units to enable multicomponent tasks
- The tool is able to check a student’s answer on equivalence through a computer algebra engine
- The tool enables the student to use a computer algebra system as a tool
- The tool is able to display formulas correctly.

*Tool category (contains criteria related to tool use)*

- The tool has an authoring function that enables teachers to add or modify content. (e.g. questions, texts, links, graphs, feedback)
- The tool is easy to use for a student (e.g. equation editor, short learning curve, interface)
- The tool is accessible anytime, anywhere
• The tool stores the answers given by a student

• The tool stores the solution process of the student

• The tool provides the author/teacher with question management facilities

• The tool makes use of standards (e.g. QTI, SCORM)

• The tool has readily available content

Assessment category (contains criteria related to assessment)

• The tool provides several assessment modes (e.g. practice, test)

• The tool caters for several types of feedback (e.g. conceptual, procedural, corrective)

• The tool takes the mastery and profile of the student into account and serves up appropriate questions (adaptivity).

• The tool has a review mode showing what the student has done wrong or right

• The tool allows for the use of several question types (e.g. multiple choice, open)

General category (contains general criteria for a digital tool for education)

• The cost of the tool

• The licensing of the tool (e.g. open, proprietary)

• The technical support of the tool

• The languages the tool supports

• The stability and performance of the tool

• The structure and attractiveness of the tool’s interface
Focus on High School Mathematics: Technology to Support Reasoning and Sense Making Criteria

- Is the action purposeful, thoughtful, and deliberate?
- Are any special technical skills needed?
- Is the action mathematically meaningful?
- Are the consequences immediate, visual, and mathematically meaningful?
- Is the connection between the action and consequence evident?
- Does the use of technology provide enhancements?
- Are the representations mathematically faithful (true to the math) and cognitively faithful (not perceptually misleading)?
- Does the action-consequence scenario provide opportunities for reflection?
- What prior mathematics knowledge and understandings will students need?
- What new knowledge and understandings could students build?
# APPENDIX K

## The Newly Created Evaluation Instrument

<table>
<thead>
<tr>
<th>Name of app: ____________________________________________</th>
<th>Criteria not met</th>
<th>Criteria met well</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
</tbody>
</table>

### 1. Meaningfully interactive
The user is an active participant. He/she has a choice of actions, and these actions shape and change the information being displayed.

### 2. Natural actions on objects
Actions are performed by gestures with the objects. Gestures are congruent to the mathematical action being performed, resulting in objects behaving as expected.

### 3. Visible consequences of actions
The consequences of the user's actions are visible. Changes that happen are expected and noticeable.

### 4. Range of actions not limited by technology
The app takes advantage of the medium, allowing the user to perform any mathematical action that they could with pen-and-paper, and possibly more.

### 5. Necessitate thoughtful, deliberate, actions
Thoughtful and deliberate actions are required, not random ones. The user is unlikely to be able to progress without carefully considering his or her actions.

### 6. Cognitively faithful interactions
Through using the app, the user will better understand the mathematical concepts and objects within. How and why mathematical procedures work are explained.

**Interactions category score** (average score of items 1-6)

### 7. Mathematically faithful
Mathematical content is accurate and error-free.

### 8. Cognitively demanding
Requires higher-level thinking skills, rather than simple memorization or application of procedures. The user is expected to synthesize previous knowledge with new information to solve problems in non-procedural ways.

### 9. High-quality presentation
Multimedia (graphics, audio, animations, etc.) are incorporated in ways that contribute to the user's learning.

**Quality of content category score** (average score of items 7-9)
<table>
<thead>
<tr>
<th>Item</th>
<th>Criteria not met</th>
<th>Criteria met well</th>
</tr>
</thead>
<tbody>
<tr>
<td>10. Provides relevant and timely feedback</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feedback (either explicit or implicit) is provided that helps the student progress in the task and helps the student in their understanding of the mathematical content.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Provides scaffolding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explicit scaffolding (e.g., hints, guiding questions) and implicit scaffolding (e.g., gradual increase of difficulty, building upon concepts) keeps the user from getting stuck.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Provides opportunity for reflection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Promotes and stimulates reflection, possibly with a history of actions, sense-making questions, or deliberate pauses at key moments.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feedback and support category score (average score of items 10-12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Easy to use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usable without much documentation, technical skills, or instruction. “Pick-up-and-play”.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Documentation is clear and easy to understand</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Documentation and help files are present if they are needed. Help files include pictures, animations, or examples. Help files are easy to navigate.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Interface is clear, consistent, and follows technology standards</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buttons clearly identified and interacted with. Text is large enough to be readable. No unnecessary clutter. Pages have a consistent look. Standard icons are used for common actions.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Usable by a variety of learners</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accessible to users of various abilities, learning styles and backgrounds. Allows multiple entry points (e.g. difficulty levels), multiple approaches, and multiple solution strategies.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ease of use category score (average score of items 13-16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall score (average of the four category scores)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>