Full length article

An examination of children's learning progression shifts while using touch screen virtual manipulative mathematics apps

Christina M. Watts [a], Patricia S. Moyer-Packenham [a, *], Stephen I. Tucker [b], Emma P. Bullock [a], Jessica F. Shumway [a], Arla Westenskow [a], Jennifer Boyer-Thurgood [a], Katie Anderson-Pence [c], Salif Mahamane [a], Kerry Jordan [a]

[a] The Virtual Manipulatives Research Group at Utah State University, 2605 Old Main Hill, Logan, UT, 84322, USA
[b] Virginia Commonwealth University, 821 W. Franklin Street, Richmond, VA, 23284, USA
[c] University of Colorado Colorado Springs, Columbine 3045, 1420 Austin Bluffs Pkwy, Colorado Springs, CO, 80918, USA

ABSTRACT

The purpose of this study was to examine shifts in young children's learning progression levels while they interacted with virtual manipulative mathematics apps on touch-screen devices. A total of 100 children participated in six mathematics learning sequences while using 18 virtual manipulative mathematics touch-screen apps during clinical interviews. Researchers developed a micro-scoring tool to analyze video data from two camera sources (i.e., GoPro camera, wall-mounted camera). Our results showed that it is possible to document evidence of shifts in children's learning progressions while they are interacting with mathematics apps on touch-screen devices. Our results also indicated patterns in the children's interactions that were related to the shifts in their learning progression levels. These results suggest that an open-ended number of tasks with a variety of representations and tasks at varying levels of difficulty led to children refining their understanding and shaping their concept image of mathematical ideas resulting in incremental shifts in learning. The results of this study have important implications about how mathematical tasks in touch-screen apps may prompt children's incremental learning progression shifts to occur, and thereby promote opportunities for learning. We propose that design features in mathematics apps can be created to support and encourage these learning shifts.

Published by Elsevier Ltd.

1. Introduction

Mathematical learning is a dynamic process. A child's knowledge of mathematics evolves as the child learns and grows in mathematical understanding. The Pirie and Kieren (1994) model for the growth of mathematical understanding emphasizes that learning has shifts to the inner and outer levels of their model for any given mathematical topic that is learned. Terminology, such as *learning progressions* and *developmental progressions*, have been used to describe the developmental and fluid aspects of children's

mathematical learning. For example, Clements and Sarama (2010) describe *developmental progressions*:

> Children follow natural developmental progressions in learning and development. As a simple example, children first learn to crawl, which is followed by walking, running, skipping, and jumping with increased speed and dexterity. Similarly, they follow natural developmental progressions in learning math; they learn mathematical ideas and skills in their own way. When educators understand these developmental progressions, and sequence activities based on them, they can build mathematically enriched learning environments that are developmentally appropriate and effective. These developmental paths are a main component of a *learning trajectory*. (p. 1)

Both learning progressions and developmental progressions have been terms used to describe the hierarchical levels that document children's understanding as a shifting process with

incremental steps that lead from limited understanding to greater understanding. It is this idea, that learning and development happens in small incremental shifts, which led to the investigation in this paper.

Sarama, Clements, Barrett, Van Dine and McDonel (2011) state that "a critical mass of ideas from each level must be constructed before thinking…becomes ascendant in the child's mental actions and behavior" (p. 668). Much like these Sarama et al. (2011), we believe that a critical mass of ideas must be constructed before children are able to demonstrate learning for a particular mathematics concept, and that the incremental shifts in the development of these ideas can be documented by examining children's learning progressions. We hypothesize that an examination of shifts in children's learning progressions may lead to a deeper understanding of how those shifts lead to learning when children use virtual manipulative mathematics apps on touch-screen devices. To examine this hypothesis, we conducted interviews with young children, asking questions on several mathematics topics. During the interviews, children used a variety of virtual manipulative mathematics apps on touch-screen devices to learn about the mathematics topics. By conducting an in-depth analysis of individual children and documenting the incremental shifts in their developing mathematical ideas, this paper contributes important insights about what children were doing when shifts in their learning progressions occurred. The results of this study have important implications about how mathematical tasks in touch-screen apps may prompt children's incremental learning progression shifts to occur, and thereby promote opportunities for learning. We propose that design features in mathematics apps can be created to support and encourage learning shifts.

2. Theoretical framework

2.1. Learning progressions

In this paper we chose to use the term learning progression (rather than developmental progression) to refer to observable changes in knowledge and mathematical skills. Winick, Avallone, Smith, and Crovo (2008) define a learning progression as "a sequence of successively more complex ways of reasoning about a set of ideas" that can follow multiple interconnected pathways (2008, p. 90). Smith, Wiser, Anderson, and Krajcik (2006) similarly define learning progressions as "descriptions of successively more sophisticated ways of reasoning within a content domain based on research syntheses and conceptual analyses" (p. 2). The learning progression concept, as proposed by Smith et al., is embedded in Clements and Sarama's (2004) term "learning trajectory" which they define as "descriptions of children's thinking as they learn to achieve specific goals in a mathematical domain, and a related, conjectured route through a set of instructional tasks designed to engender those mental processes or actions hypothesized to move children through a developmental progression of levels of thinking" (p. 83). Learning progressions have been used to analyze students' understanding of a wide variety of subjects, such as matter and atomic molecular theory (Smith et al., 2006), modern genetics (Duncan, Rogat, & Yarden, 2009), scientific argumentation (Berland & McNeill, 2010), energy concepts (Lee & Liu, 2010), celestial motion (Plummer & Krajcik, 2010) and the nature of matter (Stevens, Delgado, & Krajcik, 2010). These examples of learning progression applications describe changes in children's understanding over a long period of time. Learning progressions can also be used to describe changes in children's understanding over a short period of time. As evidenced by the applications above, the long term type of learning progression is becoming more prevalent. However, learning progressions which describe change in a short

period of time are still uncommon in the research literature.

Because of the emphasis on learning mathematics as a process, and that children's learning follows a learning progression (or developmental progression), some researchers reject the notion that a child is either "correct" or "incorrect" in her thinking when doing mathematics. Researchers who support the idea of learning progressions and constructivist theories of learning suggest that children are not "correct" or "incorrect", but rather, they are somewhere on a continuum between a more primitive understanding of the mathematical topic and a more complete understanding of the mathematical topic (Pirie & Kieren, 1994). Constructivist theories of learning suggest that cognitive structures do not progress evenly, and that cognitive structures of children may experience forwards and backwards discrete shifts in understanding along a continuum. Because of this continuum toward understanding, as children restructure their thinking, there is an observable restructuring in their performance (Piaget, 1946/1970).

Stroup and Wilensky (2000) suggest a framework that focuses on "…understandings that come into being in relation to activity. These understandings are constructed in ways that cannot be reduced to the individual 'responses' or 'contingencies' of performance, or to the linear summation (accumulation) of these 'responses' or 'contingencies'" (p. 900). In this type of framework, Stroup and Wilensky (2000) state that "The movement of groups of individuals along the range of values (scores on a scale of performance) will not be smooth. Instead, the movement will be characterized by more or less discontinuous jumps between the modes associated with the activity of certain structures or ways of understanding" (p. 902). In alignment with our hypothesis that an examination of shifts in children's learning progressions may lead to a deeper understanding of how those shift lead to learning, this study did not seek out large increases in scores for these brief interviews, nor were we looking for correct versus incorrect answers. Instead, we examined where each child was on a continuum of learning for each mathematical topic at the beginning of the interviews, and focused our attention on documenting shifts in children's learning progression levels as they occurred while children were interacting with virtual manipulatives mathematical apps.

2.2. Mathematics learning and technology

The National Council for Teachers of Mathematics (NCTM, 2000) includes the use of technology as one of their main principles for teaching school mathematics. They state that "technology is essential in teaching and learning mathematics; it influences the mathematics that is taught and enhances student learning" (p. 24). For this reason, the choice of touch-screen virtual manipulative mathematics apps versus mouse-driven apps was intentional. Segal (2011) found that:

It appears that the touch interface allowed a better flow of interaction. This is a behavioral mapping property that allows children better control of the interaction, and reduces the mental effort required by working memory. It supports the findings of Revelle and Strommen (1990) with respect to the ease of use of a touch screen for younger children, compared to a mouse-based interface. (p. 36)

Current research in mathematics using touch-screen devices has shown positive student outcomes. For example, Spencer (2013) reported significant gains in number recognition and digit formation for children aged 4–5. Similarly, Riconscente (2012, 2013) reported gains in fraction ability for fifth-graders when touch-screen devices were used. The variety of studies in which children used

touch-screen devices has recently been growing (Bartoschek, Schwering, Li, Münzer, 2013; Bertolo, Dinet, & Vivian, 2014; Chen, 2011, 2012; Haydon et al., 2012; Kilic, 2013; Ladel & Kortenkamp, 2013; Reid & Ostashewski, 2011). However, very few of these studies focus on the small shifts in understanding that can occur when children are using touch-screen devices.

In addition to the choice to use touch-screen mathematics apps for this study, we also selected apps that were classified as *virtual manipulatives*. Virtual manipulatives have been defined by Moyer, Bolyard, and Spikell (2002) as "an interactive, Web-based visual representation of a dynamic object that presents opportunities for constructing mathematical knowledge" (p. 373) and more recently defined by Moyer-Packenham and Bolyard (2016) as "an interactive, technology-enabled visual representation of a dynamic mathematical object, including all of the programmable features that allow it to be manipulated, that presents opportunities for constructing mathematical knowledge" (p. 5). The use of virtual manipulatives for mathematics learning is well supported by representation theory (Goldin & Kaput, 1996; Goldin, 2003) and how interactions with internal and external representations lead to learning (Manches & O'Malley, 2012). Moyer-Packenham and Westenskow's (2013) meta-analysis using 32 research reports and 82 effect size scores revealed overall moderate effects (0.34) on student achievement when virtual manipulatives were compared with other instructional treatments. In addition, this meta-analysis found that there were five categories of virtual manipulative apps linked to positive learning outcomes:

> *focused constraint* (i.e., VMs focus and constrain student attention on mathematical objects and processes), *creative variation* (i.e., VMs encourage creativity and increase the variety of students' solutions), *simultaneous linking* (i.e., VMs simultaneously link representations with each other and with students' actions), *efficient precision* (i.e., VMs contain precise representations allowing accurate and efficient use), and *motivation* (i.e., VMs motivate students to persist at mathematical tasks) (p. 35).

A final important consideration for this research was the use of videos to capture and interpret children's learning progressions during the interviews. In a review of 45 educational applications for the iPad, Falloon (2013) reported that some features of apps intended to aid children, such as links to relevant websites, caused confusion that actually hindered children's progress towards the learning objectives of the apps. Additionally, Dunleavy, Dede, and Mitchell (2009) found that students became overwhelmed by a game feature that was originally intended to enhance their experience. These unintended consequences are often revealed only through a close examination of a child's interactions with the app. The precision of video to record and play back children's actions allows a focused opportunity to observe and document these interactions, and thereby capture and reveal small shifts in children's learning progression levels during their mathematics app interactions. Understanding what is happening when there are shifts in children's learning is just as important as understanding what is happening when there are not any shifts in children's learning. For that reason, we took a balanced approach to our data analysis that included both statistical and graphical analyses to examine changes in children's learning progression levels.

3. Research questions

Our overarching hypothesis was that examining shifts in children's learning progressions may lead to a deeper understanding of how those shifts lead to learning with virtual manipulative mathematics apps on touch-screen devices. The following research questions were the focus of this inquiry: What is the evidence of shifts in children's learning progressions when young children use virtual manipulative mathematics apps on touch-screen devices? How do patterns in the children's activities with the mathematics apps relate to shifts in children's learning progressions?

4. Methods

The overarching research design for this study was a convergent mixed methods design (Creswell & Plano Clark, 2011; Tashakkori & Teddlie, 2010). For this design, we collected primary quantitative and qualitative data to answer complementary research questions, and then merged the data following analysis, to address our overarching hypothesis. The rationale for this design is that the researcher wants to obtain complementary data on the same topic to better understand the research problem and then examine the complementary sets of results together to allow an overall interpretation. This type of design "lends itself to team research, in which the team can include individuals with both quantitative and qualitative expertise" (Creswell & Plano Clark, 2011, p. 78).

4.1. Participants

Researchers recruited 100 children from public and charter elementary schools, a university preschool, and a university laboratory school using information brochures. We recruited children in three categories: Preschool, ages 3—4, ($N = 35$); Kindergarten, ages 5—6, ($N = 33$); and Grade 2, ages 7—8, ($N = 32$). Most children were Caucasian (89%), and 34% reported low socioeconomic status. Parents reported on children's use of touch-screen devices (TSDs). Eleven percent of children had more than five TSDs, 78% had between one and four, and 8% had none. In addition, 13% of children had access to their own TSD at home. Forty-five percent of children used the TSDs every day, 2% used it 4—5 days per week, 40% used it 1—3 days per week, 10% never used it, and 3% did not respond.

4.2. Procedures

There were 18 virtual manipulative apps selected for the study. Each child interacted with six apps. Each set of three apps formed a learning sequence that focused on a specific mathematics topic, with two mathematics topics for each age group. The apps were selected based on Ginsburg, Jamalian, and Creighan's (2013) cognitive principles for app design (e.g., mathematically appropriate content, use of appropriate models, and appropriate physical interactions) and the criteria for high quality virtual manipulatives with affordances shown to impact student learning (see Moyer-Packenham et al., 2015b). The 18 apps are presented as figures in the results section of the paper.

The children in the study participated in one-on-one clinical interviews in a single 30—40 min session in rooms equipped with two-way mirrors, audio observer booths, and built-in wall-mounted video cameras. Each child was equipped with a GoPro camera positioned to see the child's actions and interactions with the touch-screen apps, providing a secondary view (Roschelle, 2000). Interviewers gave children iPad devices with the mathematics apps and used interview protocols to engage each child with the apps. Interviews were video recorded for later description and coding. Video recording is important for capturing interactions of children with mathematics apps on touch-screen devices because many of the decisions that children make occur in quick succession and they

cannot reliably be coded and analyzed in the moment by observers. Video recordings make it possible for multiple researchers to analyze different features of the interview in repeated viewings (DeCuir-Gunby, Marshall, & McCulloch, 2012).

During the interviews, the children put on a GoPro camera positioned to view their interactions with the apps. Two researchers, who were experienced classroom teachers, conducted each interview. One researcher interviewed the child while the other researcher took notes from the observation booth. The interviewer followed the established interview sequence which included two different learning sequences of tasks, each with a Pre App, two Learning Apps, and a Post App (see Appendix A).

4.3. Video coding: developing tools to measure learning progression levels

Adesina, Stone, Batmaz, and Jones (2014) use the word "trace" to describe following "the path or history" of a child's learning and interactions within an app. Using video recordings is one way to have a record that allows us to trace each child's learning progression levels throughout the learning sequences. To trace the path of children's movement along the continuum of the learning progression for each mathematical topic, we created micro-scoring tools that could identify small shifts along a learning progression continuum for each app. The micro-scoring tools were developed to code the videos for four variables: learning progression scores on the Pre App assessment, the Post App assessment, the Learning App1 assessment, and the Learning App2 assessment. To establish validity and reliability of the micro-scoring tools for a mixed method study, Creswell and Plano Clark (2011) recommend using rigorous practices of checking the quality and the interpretation of the data, such as data triangulation (Merriam, 2009; Yin, 2009). We achieved data triangulation by examining and viewing the videos multiple times and having multiple researchers view the videos independently during the development of the micro-scoring tools (Schubert, 2009). This process involved teams of researchers collectively and independently reviewing subsets of the video data. Next, group members came together to develop the code parameters and identify examples of the observable actions and verbalizations in the video data to create a the micro-scoring tool for each app. Pairs of researchers tested the micro-scoring tools by examining 10 interviews in each age group and then refined the tools. This micro-scoring tool development process resulted in a set of 18 scoring tools for the 18 apps used in the study. To code the video data using the micro-scoring tools, we used one of the most common methods for ensuring reliability: inter-coder agreement (Creswell & Plano Clark, 2011). Groups of researchers were trained

to use the micro-scoring tools and independently coded the entire interview set (Corbin & Strauss, 2015; Saldaña, 2013; Stebbins, 2001). To ensure reliable and accurate coding, 30–100% of the video data were double-coded by two researchers. Our research team had experience using this process previously during multiple research projects (Moyer-Packenham et al., 2014).

Table 1 shows a sample of the learning progression micro-scoring tool for the preschoolers' Base-10 Counting Sequence. The highest possible learning progression level a child could attain for this sequence was a level 6; this score was rescaled to a 0 to 1.0 scale during data analysis. The lowest learning progression level, where a child either guessed or did not respond, was a level 1. The micro-scoring tool enabled us to document children's movement along the range of the learning progression rather than simply scoring children's responses as correct or incorrect.

Members of the research team trained to use the micro-scoring tools and used the tools to examine and score the videos. Researchers assigned learning progression level scores for every task completed by every child for every app (e.g., Pre App, two Learning Apps, and Post App). Two independent researchers performed over 30% of the coding for all of the videos.

4.4. Data analysis

There were four variables used to determine changes in children's learning: learning progression scores on the Pre App assessment, the Post App assessment, the Learning App1 assessment, and the Learning App2 assessment. These variables were used in three major phases of analysis: a visual analysis, a statistical analysis, and a graphical analysis. The use of different analysis methods is typical of a convergent mixed methods design because it supports researchers in answering complementary research questions, and allows researchers to merge data following the analysis to address overarching research problems (Creswell & Plano Clark, 2011; Tashakkori & Teddlie, 2010). By having complementary data to answer the research questions, we are better able to understand the phenomenon and use the complementary results in an overall interpretation. In the visual analysis, we examined the learning progression levels for all children in individual plots, which we then aggregated into histograms for the graphical analysis. In the statistical analysis, we compared learning progression levels in various apps to one another. Initially, for each child, we examined the child's attained learning progression levels for every learning task in every app. Some apps had only one task while other apps had an open-ended number of tasks. When there was more than one task, we computed averages for the learning progression

Table 1
Learning progression for preschool base-10 counting sequence app.

Level	Description of mathematical learning progression expectations
1	Child guesses; no response
2	Moving blocks as the app counts: child knows to move blocks to build an amount, but does not count aloud or exhibit cardinality.
3	Pre-counting: child says number names but does not match to objects (does not have one-to-one correspondence).
4	One-to-one correspondence (for at least three objects): child says the standard list of counting words in order and matches each spoken number with one and only one object, but cannot tell how many (e.g., does not stop at target number; is cued to stop by sparkles feedback).
5	Counting out a collection up to five: child has developed understanding of cardinality; child can count the items in a set to five and knows that the last number counted tells the size of that set (e.g., stops at target number before sparkles feedback).
6	Counting out a collection from six to ten: child has developed understanding of cardinality; child can count the items in a set to ten and knows that the last number counted tells the size of that set (e.g., stops at target number before sparkles feedback).

levels for each child's performance on every app: the Pre App, the two Learning Apps, and the Post App. For the statistical analysis, one number represented the child's average learning progression level on multiple tasks within each app experience. We also examined the forwards and backwards shifts in children's attained learning progression levels within each app and quantified this using the variance. There were three aspects of each child's interview used in the statistical comparisons: the variance of a child's scores on the tasks within a learning progression, the average learning progression level a child attained on each individual app, and (when there was an open-ended number of tasks) the number of tasks the child completed with each app.

Using the average learning progression levels on the Pre and Post Apps, we examined the overall change in each child's learning progression levels between the two apps. This helped us to identify children whose learning progression levels increased, stayed constant, or decreased during the learning sequences. We used this information to examine relationships between children's performance on the two Learning Apps and the Pre and Post Apps. Pearson's r correlations were used analyze these relationships, rating as a no to negligible correlation for r between 0 and ± 0.19, weak correlation for r between ±0.20 and ± 0.29, moderate correlation for r between ±0.30 and ± 0.39, strong correlation for r between ±0.40 and ± 0.69, and very strong correlation for r greater than ±0.70 (Cohen, 2013). During the graphical part of the analysis, we created histograms of children's learning progression levels in each app to focus on Pre to Post App learning shifts for each age and created histograms to show examples of these shifts.

5. Results

The results below include a description of each learning sequence with a sample child's interview experience from the visual analysis, and results of the statistical and graphical analysis. In doing so, we present the data so that small shifts in children's learning progressions are visible through the evidence.

5.1. Preschool, ages 3-4

The 35 preschool children worked with two different learning sequences. The first learning sequence, Base-10 Counting, used the *Montessori Numbers (Quantity: 1–9)*, *Montessori Numbers (1 to 20: 1–5)*, and *Montessori Numbers (Numerals from Quantity: 1–9)*. Fig. 1 shows a typical preschooler's interview experience in this learning sequence from the visual analysis. The second learning sequence, Seriation, used *Intro to Math* (*Red Rods*), *Pink Tower (Free Moving)*, and *Pink Tower (Tap)*. Fig. 3 shows a typical preschooler's interview experience in this learning sequence.

5.1.1. Preschool Base-10 Counting Learning Sequence

In the Preschool Base-10 Counting Learning Sequence, the Pre App required children to use blocks to build numbers between 1 and 5 and then build numbers between 6 and 9. In Learning App 1, children see a demonstration of building the numbers 1−5 and are then asked to construct the numbers using blocks. In Learning App 2, children count the number of blocks that are generated and then choose the numeral associated with the quantity. Children completed an open-ended number of tasks in the 4 min allowed. In the Post App children built numbers between 1 and 9 using the blocks.

Fig. 1 shows screenshots of the apps with a sample child's learning progression levels for each app. In the Pre App this child attained a 1.0, the highest possible score, on the first task and a 0.6 on the second task. In Learning App 1, this child attained 1.0's on all three tasks. In Learning App 2, this child completed nine counting tasks in the allotted time and performance on these tasks shifted between learning progression levels of 0.5 and 0.9. On the Post App, the child obtained a learning progression level of 1.0 on both tasks.
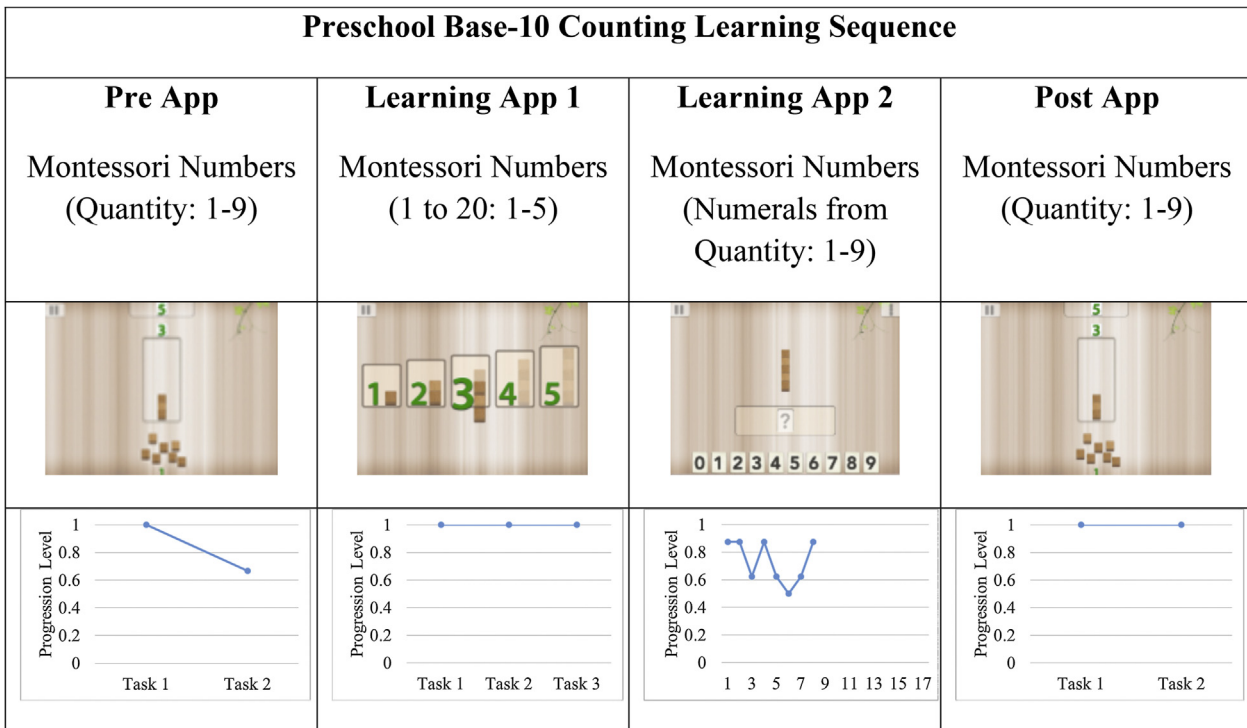

Fig. 1. Sample preschool child's learning progression on Base-10 Counting Sequence. A score of 1.0 indicates the highest level of performance on the learning progression for each app.
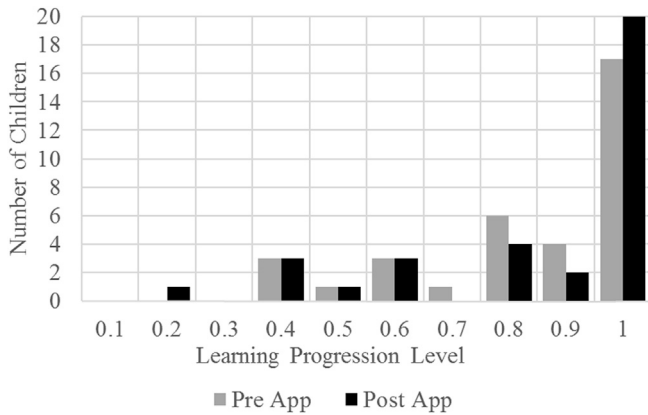
Fig. 2. A histogram of the Preschool Base-10 Counting Learning Sequence learning progression levels from the Pre App and Post App. Each vertical unit represents an individual child's average. A score of 1.0 indicates the highest learning progression level.

A statistical analysis of the learning progression levels for all preschool children on the apps in the Preschool Base-10 Counting Learning Sequence showed that the children's learning progression level on Learning App 1 tasks and Learning App 2 tasks had a strong correlation ($r(35) = 0.40$, $p = 0.018$). This means that those children who performed well on Learning App 1 also performed well on Learning App 2. For Learning App 2, with an open-ended number of tasks, there was a strong correlation ($r(35) = 0.41$, $p = 0.013$) between children's learning progression level on the tasks and the number of tasks they completed. In other words, the more tasks children completed, the higher the learning progression levels the children attained on Learning App 2. This shows that children who completed a large number of tasks were not just rushing through

with incorrect answers or guesses, they were successfully completing the tasks. One child completed twice as many tasks as the other children, potentially skewing the correlation. Excluding this child, the correlation was still moderate ($r(35) = 0.34$, $p = 0.045$). There were no other significant correlations for this sequence.

In the graphical analysis, we can see all children were at or above 0.4 for their learning progression level on the Pre App (see Fig. 2). Thirty-two of the 35 (91.4%) children increased or stayed constant from the Pre App to the Post App. This graphical representation shows that there were small shifts in learning for most of the children. There were three children who attained lower learning progression levels between the Pre and Post Apps. This seemed to be influenced by a feature of the app − celebration sparkles, which were intended to indicate a correct response. However, the amount of time between the construction of the correct number and the deployment of the sparkles prompted three children to continue adding blocks to the construction. Consequently, the celebration sparkles aligned with an incorrect construction.

5.1.2. Preschool Seriation Learning Sequence

In the Preschool Seriation Learning Sequence, children began with a Pre App, creating a tower of pink blocks from largest to smallest. Children then completed Learning App 1, which provided a more scaffolded experience building a pink tower. Learning App 2 required children to order rods from longest to shortest. On the Post App, children repeated the tower building task.

Fig. 3 shows apps in the learning sequence with a screen shot of one child's experience and a sample of their learning progression levels for each task. This child began with a 0.4 learning progression level on the single task in the Pre App and attained a 0.8 learning progression level on the single task in the Post App. Since there is only one task in each of these apps the learning progression level is

Fig. 3. Sample child's learning progression in the Preschool Seriation Learning Sequence. A score of 1.0 indicates the highest level of performance on the learning progression for each app.
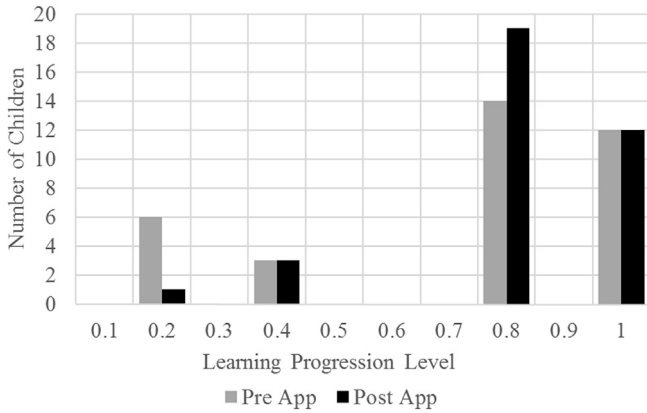
Fig. 4. A histogram of the Preschool Seriation Learning Sequence learning progression levels from the Pre and Post Apps. Each vertical unit represents an individual child's learning progression level. A score of 1.0 indicates the highest learning progression level.

indicated by a single dot in the figure.

The statistical analysis of the learning progression levels for all preschoolers on each of these apps presented no variance as there was only one task per app. There were no significant correlations for performance among the apps in this sequence. The graphical analysis revealed several shifts not indicated by the statistical analysis. Fig. 4 shows a histogram of the learning progression levels of the preschool children on the Pre and Post Apps. Thirty-one of the 33 (93.9%) preschool children obtained learning progression scores that increased or stayed constant between the Pre and Post Apps. For this sequence, several children made a noticeable shift to higher learning progression levels. The main shift occurred for children who scored a 0.2 level on the Pre App. Children who were

already at the highest level of the learning progression on the Pre App maintained this level on the Post App.

5.2. Kindergarten, ages 5-6

Thirty-three kindergartners participated in two different learning sequences, the Base-10 Quantity Learning Sequence and the Subitizing Learning Sequence. Two of the kindergarten children did not fully complete the Subitizing Learning Sequence, but they did complete the Base-10 Quantity Learning Sequence. The Base-10 Quantity Learning Sequence, used *Montessori Numbers (Quantity: 10–99)* for the Pre and Post Apps and *Montessori Numbers (1 to 20: 11–20)* and *Montessori Numbers (Numerals from Quantity: 10–99)* as Learning Apps. Fig. 5 shows a typical kindergartner's interview experience in this learning sequence. The second learning sequence, Subitizing, used the *Friends of Ten (Teaching Tool)* for Pre and Post Apps with *Hungry Guppy* and *Fingu* as Learning Apps. Fig. 7 shows a typical kindergartner's interview experience in this learning sequence.

5.2.1. Kindergarten Base-10 Quantity Learning Sequence

The Kindergarten Base-10 Quantity Learning Sequence began with the Pre App where children built three different two-digit numbers using the tens rods and the single units. For Learning App 1 children observed a demonstration of two-digit numbers being built and then constructed four numbers on their own. Learning App 2 required children to count the number of blocks displayed with the tens rods and units. Here children needed to recognize that the tens rods were composed of ten units. Children chose the correct numerals associated with the two-digit number. The Post App asked children to construct three two-digit numbers.

Fig. 5 shows a sample of a child's learning progression levels for each app. In the three tasks on the Pre App this child obtained
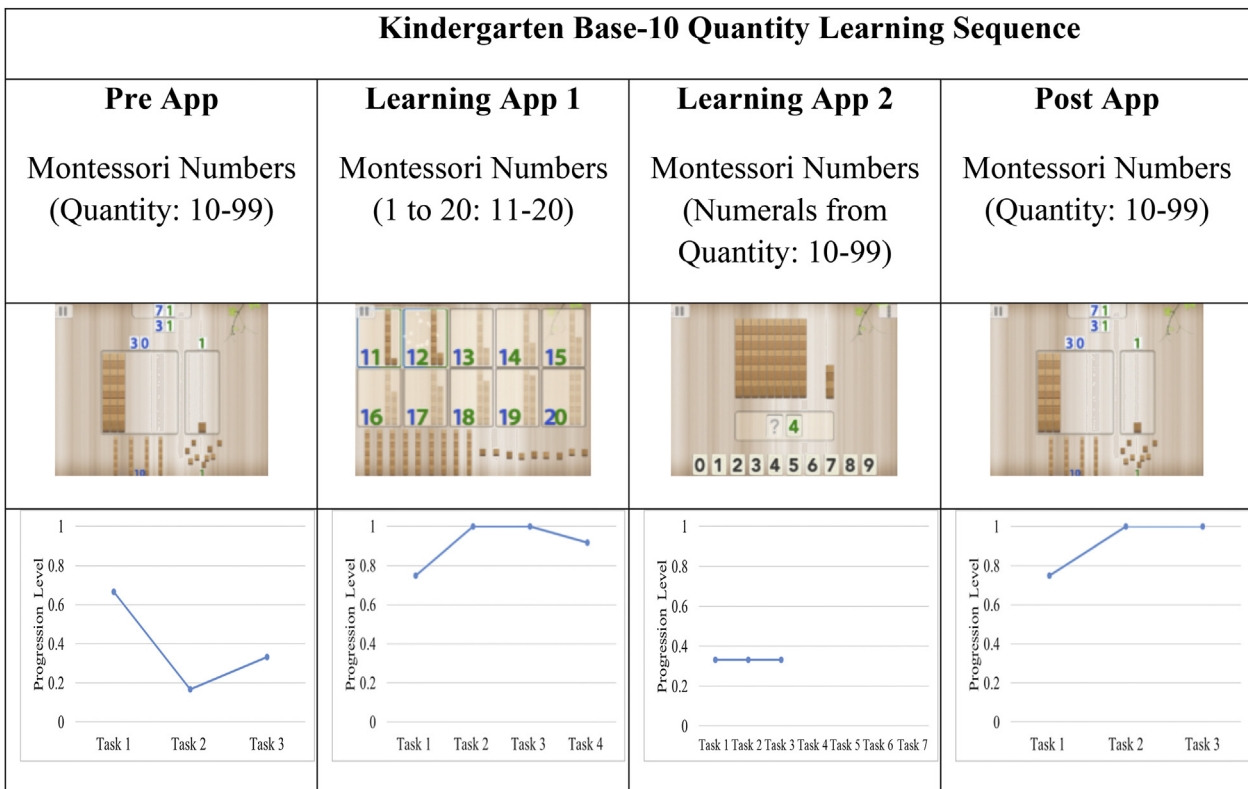

Fig. 5. Sample child's learning progression in the Kindergarten Base-10 Quantity Learning Sequence. A score of 1.0 indicates the highest level of performance on the learning progression for each app.
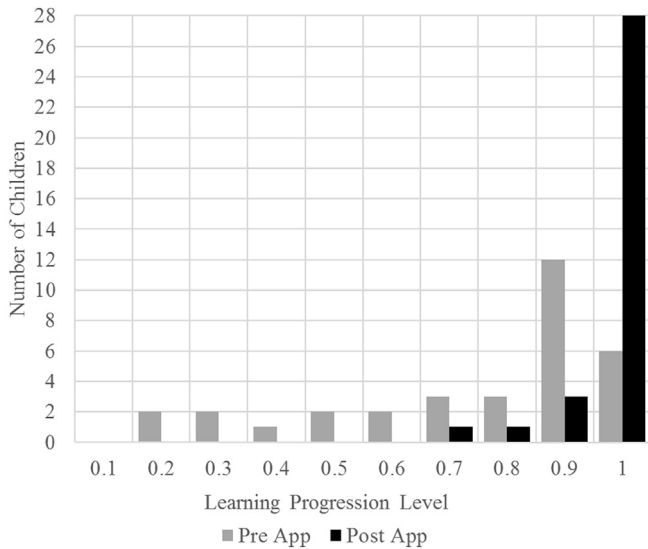
Fig. 6. Histogram of Kindergarten Base-10 Quantity Learning Sequence learning progression levels from Pre and Post Apps. Each vertical unit represents an individual child's average. A score of 1.0 indicates the highest learning progression level.

learning progression levels of 0.7, 0.1, and 0.3, respectively. In Learning App 1, which provided a scaffolded experience, the child obtained high progression levels on all four tasks. In Learning App 2, there was an open-ended number of tasks and the child completed three tasks with lower learning progression levels on each of them. On the Post App, the child improved, scoring the highest learning progression level of 1.0 on the last two tasks.

Overall, in the Base-10 Quantity Learning Sequence, kindergarten children who completed more tasks on Learning App 2 obtained higher learning progression levels on those tasks ($r(33) = 0.53$, $p = 0.002$). This is similar to results seen for the preschool learning progression levels. Children that completed a high number of tasks when there was an open-ended number of tasks showed higher levels of achievement in each of these tasks than children who completed fewer tasks. There were no other significant correlations.

Twenty-four of the 33 (72.7%) kindergartners showed an increase or stayed constant on their learning progression levels from the Pre to Post Apps. The graphical analysis shows a wide, left-skewed spread of learning progression levels on the Pre App ranging from 0.2 to 1.0 (see Fig. 6). The Post App learning progression levels show a marked improvement and a clear visualization of this shift, with the majority of children at the highest learning progression level of 1.0, and none below 0.7. The two Learning Apps may have caused this large shift due to the alignment of the tasks among the apps.

5.2.2. Kindergarten Subitizing Learning Sequence

The Kindergarten Subitizing Learning Sequence began with the Pre App where children completed three tasks to determine the number of chips to add to create the target number. In Learning App 1 children combined numbers of dots to add up to the target number and this app had an open-ended number of tasks. Learning App 2 required children to use the correct finger placement on the screen to correspond to the number of dots displayed. This app also had an open-ended number of tasks. Children completed three tasks on the Post App.

Fig. 7 shows screen shots of each app with a sample of one child's learning progression levels on each task. This child obtained
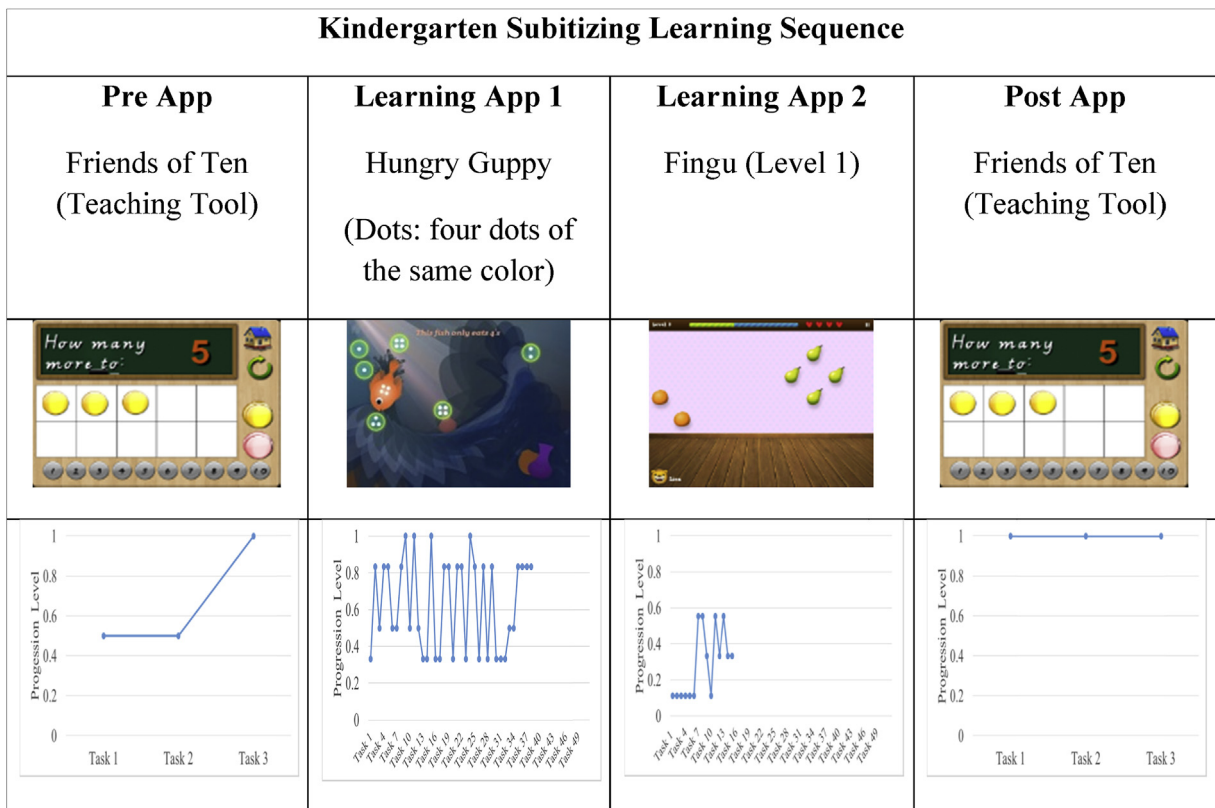

Fig. 7. Sample child's learning progression in the Kindergarten Subitizing Learning Sequence. A score of 1.0 indicates the highest level of performance on the learning progression for each app.

learning progression levels of 0.5 on the first two tasks of the Pre App and then 1.0 on the last task. This child completed 40 tasks on Learning App 1 and 15 tasks on Learning App 2. The two Learning Apps demonstrate how the child's performance varied across many tasks. On the Post App, the child improved overall with learning progression levels of 1.0 on all three tasks.

Overall, the statistical analysis for kindergarteners on the Subitizing Learning Sequence showed a moderate correlation ($r(33) = 0.37$, $p = 0.03$) between children's learning progression level changes (i.e., Pre to Post) and variability in their learning progression levels for Learning App 1. That is, children who had large or numerous shifts in their learning progression levels (e.g., obtaining 0.8, 0.3, 1.0, and 0.2) on Learning App 1 were more likely to show improvement between the Pre and Post Apps. There was also a strong negative correlation ($r(33) = -0.44$, $p = 0.009$) between the learning progression levels in Learning App 1 and the shifts in children's performance on Learning App 2. If a child obtained low learning progression levels on Learning App 1, they were likely to exhibit more variance, or shifts, in their learning progression levels on Learning App 2. One potential cause for this is that children who initially had difficulty understanding the mathematical tasks in Learning App 1 became more proficient when they attempted Learning App 2. In addition, children who completed more tasks on Learning App 2 scored well on each task of the app ($r(33) = 0.52$, $p = 0.002$), a strong correlation. This indicates that the children who completed a large number of tasks for Learning App 2 not only completed the tasks quickly, they also completed them accurately. There were no other significant correlations.

Twenty-nine of the 31 (93.5%) children obtained learning progression levels that increased or stayed constant from the Pre to Post App on this sequence. The graphical analysis shows that the majority of kindergartners obtained learning progression levels at the highest possible level on the Pre App (see Fig. 8). On the Post App, however, only three children obtained a learning progression level below 0.9, with the majority achieving the highest level.

5.3. Grade 2, ages 7–8

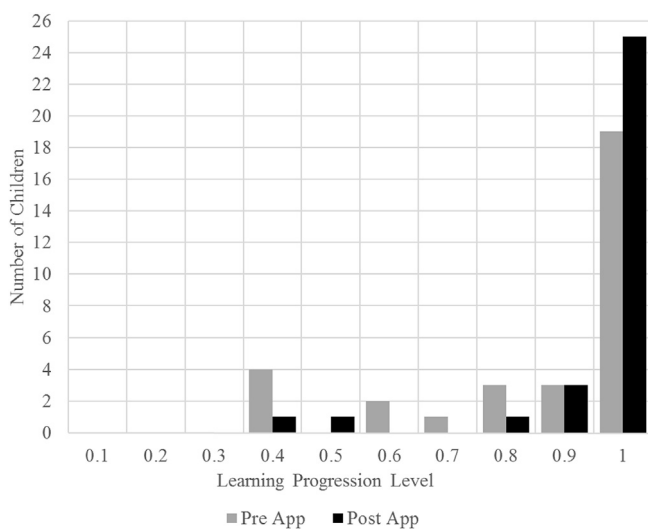Twenty-six Grade 2 children participated in two separate

Fig. 8. A histogram of the Kindergarten Subitizing Learning Sequence learning progression levels from the Pre and Post Apps. Each vertical unit represents an individual child's average. A score of 1.0 indicates the highest learning progression level.

learning sequences, the Base-10 Place Value Learning Sequence and the Skip Counting Learning Sequence. The Base-10 Place Value Learning Sequence used *Montessori Numbers (100–999)* as a Pre App. *Math Motion Zoom* and *Place Value Cards* were the two Learning Apps. *Montessori Numbers (100–999)* was the Post App. Fig. 9 shows a typical second grader's interview experience in this learning sequence. The second learning sequence, the Skip Counting Learning Sequence, used the *100s Board* as a Pre App, *Number Lines (Skip Counting)* and *Skip Counting Beads* as the two Learning Apps, and the *100s Board* as a Post App. Fig. 11 shows a typical second grader's interview experience in this learning sequence.

5.3.1. Grade 2 Base-10 Place Value Learning Sequence

The Grade 2 Base-10 Place Value Learning Sequence began with a Pre App where children built six three-digit numbers using the Base-10 blocks. Learning App 1 required the children to zoom in and out on a number line to place numbers. In Learning App 2 children built three-digit numbers using numeral cards for each place value and then constructed the final three-digit number (e.g., 995 is written as 900 + 90 + 5). Both Learning Apps had an open-ended number of tasks. The Post App asked children to construct six three-digit numbers.

Fig. 9 shows screen shots from each of the apps in the Base-10 Place Value Learning Sequence with two samples of children's learning progression levels. On the Pre App, Child 82 obtained the highest learning progression level of 1.0 on all of the tasks. On Learning App 1, this child's learning progression levels shifted forwards and backwards on the 45 tasks he completed. This child completed 20 tasks at the highest level of the progression on Learning App 2. On the Post App, this child completed all of the tasks at the highest learning progression level.

In the Grade 2 Base-10 Place Value Learning Sequence, there was a strong correlation ($r(26) = 0.58$, $p = 0.0004$) between the variance of a child's scores (the shifts forwards and backwards) on Learning App 1 and the number of tasks they completed on Learning App 2, which had an open-ended number of tasks. Thus, if a child's learning progression levels shifted on Learning App 1, the child was much more likely to complete more tasks on Learning App 2. Fig. 9 shows Child 82 had more shifts in Learning App 1 performance than Child 75. Accordingly, Child 82 completed a larger number of tasks (20 tasks) in Learning App 2 than Child 75 (4 tasks). In Learning App 1, those children who showed an increase in their learning progression levels from the Pre to Post Apps averaged 4.45 more tasks than those children who showed a decrease in learning progression levels from Pre to Post App. Similarly for Learning App 2, children with an increase in learning progression levels from Pre to Post Apps averaged 5.15 more tasks than those who showed a decrease from the Pre to Post Apps. Interestingly, children whose learning progression level stayed constant from the Pre to Post Apps averaged the same number of tasks as those who showed an increase, but had more shifts in their performance on Learning App 1. There were no other significant correlations.

Twenty-three of the 27 (85.2%) children's learning progression levels increased or stayed constant from the Pre to Post Apps. The graphical analysis for the Pre and Post App learning progression levels shows that the majority of the children initially scored very well on the Pre App and maintained these learning progression levels on the Post App (see Fig. 10). However, there are fewer children at the highest learning progression level of 1.0 on the Post App.

5.3.2. Grade 2 Skip Counting Learning Sequence

In the Grade 2 Skip Counting Learning Sequence, the Pre App
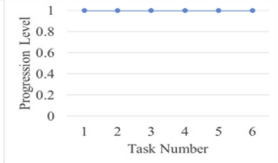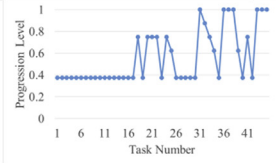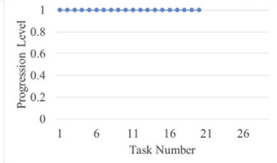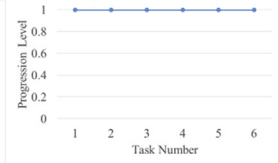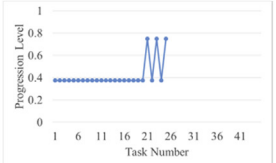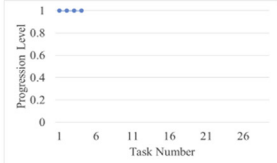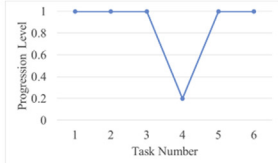
Fig. 9. Samples of children's learning progressions in the Grade 2 Base-10 Place Value Learning Sequence. A score of 1.0 indicates the highest level of performance on the learning progression for each app.
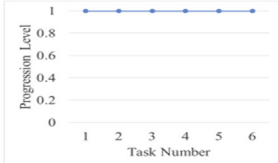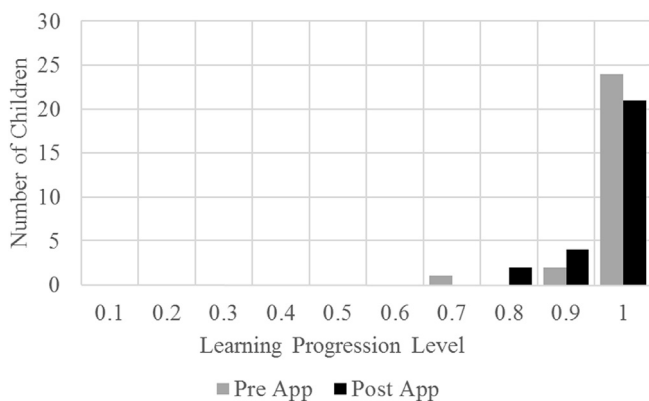

Fig. 10. A histogram of the Grade 2 Base-10 Place Value Learning Sequence learning progression levels of the Pre and Post App. Each vertical unit represents an individual child's average. A score of 1.0 indicates the highest learning progression level.

asked children to count by fours, sixes, and nines. The two Learning Apps scaffolded skip counting by providing feedback or by giving a prescribed amount to count. Learning App 1 had a fixed set of tasks; Learning App 2 had an open-ended number of tasks. They counted by fours, sixes, and nines on the Post App.

Fig. 11 shows a sample of a child's learning progression levels on the Skip Counting Learning Sequence. On the Pre App the child obtained three scores around 0.5 for skip counting by fours, sixes, and nines. On Learning App 1, the child attained the highest learning progression levels on four of the six tasks. Learning App 2 allowed for an open-ended number of tasks and the child completed ten tasks with varying learning progression levels. On the Post App, the child did not improve on the first task (skip counting by fours) but obtained the highest learning progression levels on the second and third tasks (skip counting by sixes and nines). This child remembered that there were patterns that could be followed when completing the last two tasks.

The Grade 2 Skip Counting Learning Sequence showed a moderate positive correlation of ($r(26) = 0.37$, $p = 0.037$) between the number of tasks children completed and the shifts, as measured by variance, of children's learning progression levels on Learning App 1. Again, this showed that more shifts in a child's learning progression levels correlated with a larger number of completed tasks. No other correlations were significant.

In the graphical analysis, the histogram shows the children's Pre App learning progression levels have a right-skewed distribution, but the Post App learning progression levels have a left-skewed distribution (see Fig. 12). This indicates a whole-group shift
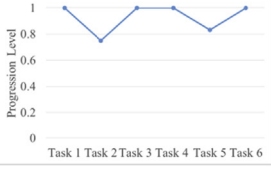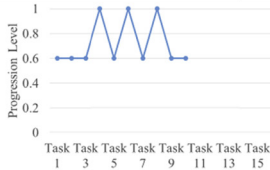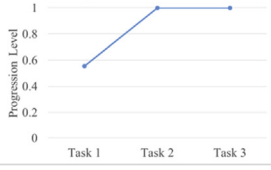
Grade 2 Skip Counting Learning Sequence			
Pre App 100s Board	**Learning App 1** Number Lines (Skip Counting)	**Learning App 2** Skip Counting Beads	**Post App** 100s Board

Fig. 11. Sample child's learning progressions in the Grade 2 Skip Counting Learning Sequence. A score of 1.0 indicates the highest level of performance on the learning progression for each app.
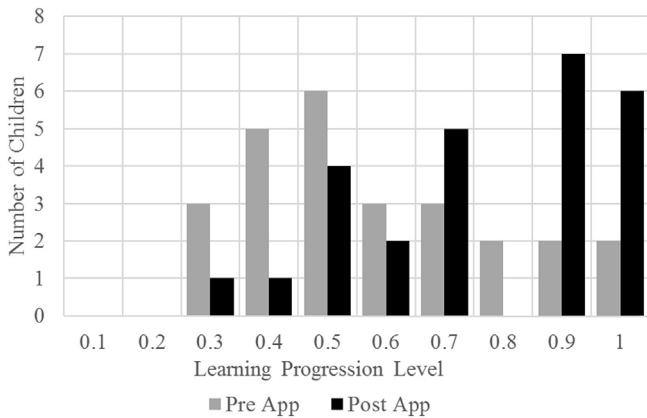
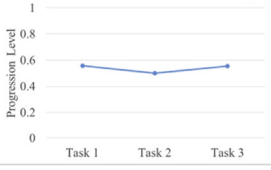**Fig. 12.** A histogram of the Grade 2 Skip Counting Learning Sequence learning progression levels from the Pre and Post App. Each vertical unit represents an individual child's average. A score of 1.0 indicates the highest learning progression level.

resulting from the children's progress through the learning sequence. All (100%) of the children obtained higher learning progressions levels on the Post App than the Pre App or stayed constant.

6. Discussion

In this study, we hypothesized that examining shifts in children's learning progression levels while they interacted with virtual manipulative mathematics apps on touch-screen devices had the potential to lead to greater understanding about how those shifts related to children's learning. Our results showed that it is possible to document evidence of shifts in children's learning progressions while they are interacting with mathematics apps on touch-screen devices. Our results also indicated patterns in the children's interactions that were related to the shifts in their learning progression levels.

6.1. Incremental shifts in learning

The visual and graphical analyses revealed incremental shifts in children's learning during their interactions with the virtual manipulative mathematics apps. Our results showed several examples of these types of shifts for individual students, for small groups of students, and for the whole group. For example, in the Kindergarten Base-10 Quantity Learning Sequence, we observed the sample child's visual analysis in Fig. 5, which showed an individual child's learning shifts. The histogram of the Preschool Seriation Learning Sequence (see Fig. 4) showed how five children made, what Stroup and Wilensky (2000) call, a discontinuous jump in their learning progression levels from the Pre App (0.2) to the Post App (0.8). This type of shift was also seen for the whole group in the Grade 2 Skip Counting Learning Sequence, where the graphical analysis showed a distinct shift from a right skew to a left skew, indicating lower learning progression levels in the Pre App and higher learning progression levels in the Post App (see Fig. 12). Documenting these discontinuous jumps allowed us to trace each child's movement along the learning progressions within each app. We would not have observed this continuum of growth without the use of a micro-scoring tool and an examination of every task every child completed throughout the learning sequences.

Learning progressions can reveal the small incremental shifts in children's cognitive structures. Clements and Sarama (2010) note that examining this movement allows researchers to answer questions such as what new objectives should be established for a particular child and what developmental tasks would

be appropriate for that child to achieve those objectives. For example, in the Grade 2 Base-10 Place Value Learning Sequence, children attained learning progression levels at or above 0.8 on the Pre and Post Apps. For this learning sequence, the initially high learning progression levels and small amount of movement indicated that these Grade 2 children were ready for an increased level of difficulty or complexity in the concept area of Base-10 Place Value Learning. Thus an overall analysis of a child's incremental shifts in learning could aid an educator in understanding a child's learning needs.

6.2. What were children doing when incremental learning shifts occurred?

The methods of data collection and analysis we used allowed us to examine children's learning on a continuum from the beginning to the end of the interview. This helped us to identify patterns in the children's interactions that were related to incremental shifts in their learning progression levels. We were able to document the forwards and backwards shifts in children's learning progression levels, especially during their interactions with the Learning Apps in each sequence. Identifying these incremental shifts between the Pre and Post Apps caused us to look more closely at what the children were doing when the learning shifts occurred.

As the results showed, there were relationships between the learning shifts and a) improvement in learning progression levels from Pre to Post Apps, b) higher attained learning progression levels in the Learning Apps, and c) a higher number of completed tasks when the app allowed for an open-ended number of tasks. When the learning shifts occurred, children were completing an open-ended number of tasks that contained multiple representations and completing tasks at varying levels of difficulty.

There were several instances where completing an open-ended number of tasks was related to improvement in children's learning progression levels between the Pre and Post Apps. For example, in the Preschool Base-10 Counting Learning Sequence and on the Kindergarten Base-10 Quantity Learning Sequence, children who completed more tasks on Learning App 2 had higher learning progression levels on those tasks. In the two Kindergarten Learning Sequences, three of the four Learning Apps had an open-ended number of tasks that children could complete. When children completed a large number of tasks they also did them well; when they completed fewer tasks they attained much lower learning progression levels.

There were several instances where completing tasks at varying levels of difficulty was related to improvement in children's learning progression levels between the Pre and Post Apps. For example, on the Kindergarten Subitizing Learning Sequence, changes in children's learning progression levels (i.e., Pre to Post) were related to the forwards and backwards shifts (as measured by variance) in children's learning progression levels on Learning App 1. On the Grade 2 Base-10 Place Value Learning Sequence, the forwards and backwards shifts in children's learning progression levels on Learning App 1 were related to the number of tasks the children completed on Learning App 2. On the Grade 2 Skip Counting Learning Sequence, the number of tasks children completed and the forwards and backwards shifts in children's learning progression levels showed that more shifts in a child's learning progression levels correlated with a larger number of completed tasks. While children were completing an open-ended number of tasks they were exhibiting forwards and backwards shifts in their performance on those tasks because the tasks had differing levels of difficulty.

6.3. How an open-ended number of tasks and varying levels of task difficulty may be related to learning shifts

Children in this study who showed numerous shifts in their learning progression levels when interacting with each app completed a greater number of tasks. Completing a greater number of tasks and completing tasks at varying levels of difficulty allowed these children to have repeated practice on the same topic with multiple representations at differing levels of challenge. Opportunities to interact repeatedly with the same mathematical topic at differing levels of challenge engages children mathematically with productive failure (Kapur, 2014), struggle, success, and persistence. Persistence and productive failure has the potential to promote a growth mindset (Dweck, 2006). In classroom application, Warshauer (2015) recommends four different teaching strategies to promote productive struggle in students, such as asking students question which refocus their thinking on concepts they have already examined or allowing students sufficient time to struggle with the tasks. Research suggests that the ability to revisit tasks is one of the key advantages of computer-based systems, such as apps in this study (Charman, 1999; Hepplestone, Holden, Irwin, Parkin, & Thorpe, 2011). With an open-ended number of tasks, children are then allowed to work with a large sample of tasks of the same type and engage in "productive failure" (Kapur, 2014), that is, learning from failed problem solving attempts and applying the knowledge to subsequent attempts. Seven of the twelve Learning Apps used in the six Learning Sequences allowed children to complete an open-ended number of tasks within an allotted amount of time. In the statistical analyses of each of these Learning Sequences, the number of tasks completed correlated with the learning progression levels on the Learning Apps, change in learning progression levels from the Pre to Post App, or shifts in learning progression levels.

One benefit of mathematics apps, such as those used in this study, is that they allow for differentiation of learners. More capable children could go faster and do more tasks and less capable students could take their time and struggle with the concepts to better understand them. Other studies have reported the importance of virtual manipulatives for students from different achievement groups. For example, Moyer-Packenham and Suh (2012) reported that students from three different achievement levels (high, average, and low) all benefited from the use of virtual manipulatives when learning fraction concepts (e.g., equivalence and addition with unlike denominators). The qualitative analysis revealed that students in different achievement groups interacted with the same virtual manipulatives in different ways and that features of the apps provided different types of scaffolds for different children's learning. The way that children access the design features of virtual manipulative apps may help or hinder children's learning. For example, Moyer-Packenham et al. (2015a) found that when young children access design features (or affordances) of apps meant to provide a helping effect, they were more likely to improve in efficiency and mathematics performance.

We theorize that these forwards and backwards shifts were a positive aspect of the children's learning during their app interactions. When children experienced an open-ended number of tasks and those tasks were at varying levels of difficulty, this gave children multiple opportunities to refine their understanding and shape their concept image (Tall & Vinner, 1981) of the mathematical idea. Having an open-ended number of tasks allowed

children to go at their own pace and to complete as many tasks as they wanted to complete. Having tasks of differing difficulty levels allowed a mixture of success and challenge during their learning experience. This process of refining their understanding and shaping their concept image appeared as forwards and backwards shifts in the children's learning progression scores, especially during their interactions with the Learning Apps. These forwards and backwards shifts during interactions with the Learning Apps were related to incremental changes between Pre and Post App learning progression levels and positive performance results.

6.4. Implications for design

We propose that design features in mathematics apps can be created to support and encourage learning shifts. Based on our results, an important implication for the design of mathematics apps on touch-screen devices is that designers must provide opportunities for an open-ended number of tasks with a variety of representations and those tasks should be mixed with varying levels difficulty. An open-ended number of tasks within an activity allows a child to refine and shape her understanding of the mathematical topic as she productively struggles in the completion of the tasks. This refining and shaping process manifests itself in the forward and backward shifts of the learning progression levels. The availability of an open-ended number of tasks with multiple representations allows children to practice the mathematical topic at differing levels of challenge while testing and refining their concept of the mathematical topic. Children's shifts in learning progression levels can be an indirect indicator of productive struggle as they work with the mathematical topic. A lack of shifting may indicate that tasks are too easy or too difficult and consequently not allowing for the refinement of understanding. Combined with other aspects of a session of learning, such as overall outcomes or the number of tasks a child can complete of a particular task type, educators can gain an in-depth as well as a big picture view of children's cognitive structures and their progression in learning.

6.5. Future research

Future research could involve designing apps based on pre-constructed learning progressions which use the feature of open-ended number of tasks with multiple representations to move the learner through a learning progression. More research is needed on the development of learning progressions as well as the effective implementation of the learning progressions within educational apps. Future research should also consider how learning progressions can be used as a formative assessment tool when coupled with technology such as educational apps. More in depth research is also needed in the development of learning progressions describing the small shifts in children's understanding which will further the goals of learning progressions which examine understanding over a longer period of time.

7. Conclusion

This study focused on documenting shifts in children's learning progressions when they used virtual manipulative mathematics apps on touch-screen devices. The results revealed patterns in children's interactions related to shifts in their learning. These results suggest that an open-ended number of tasks with a variety of representations and tasks at varying levels of difficulty may lead to children refining their understanding and shaping their concept image of mathematical ideas resulting in incremental shifts in learning. When we observed children's learning progressions shifting forwards and backwards, exhibited by variability in children's learning progression scores, this often led to positive shifts between the Pre and Post Apps. This productive failure encouraged children to test their self-constructed concept against multiple scenarios and increase to a more complete understanding of the mathematical topic.

By conducting an in-depth analysis of individual children and documenting the incremental shifts in their developing mathematical ideas, this paper contributes important insights about what children were doing when shifts in their learning progressions occurred. The results of this study have important implications about how mathematical tasks in touch-screen apps may prompt children's incremental learning progression shifts to occur, and thereby promote opportunities for learning. Design features in virtual manipulative mathematics apps, such as allowing an open-ended number of tasks with a variety of representations or varying the difficulty of tasks, can be used to support and encourage learning shifts.

Appendix A

Sequence of the interviews.

Interview	Preschool	Kindergarten	Grade 2
Pre App	Pink Tower (Free Moving)	Friends of Ten (Teaching Tool)	100s Board
Learning App 1	Pink Tower (Tap)	Hungry Guppy	Number Lines (Skip Counting)
Learning App 2	Intro to Math (Red Rods)	Fingu	Skip Counting Beads
Post App	Pink Tower (Free Moving)	Friends of Ten (Teaching Tool)	100s Board
Pre App	Montessori Numbers (Quantity: 1–9)	Montessori Numbers (Quantity: 10–99)	Montessori Numbers (Quantity: 100–999)
Learning App 1	Montessori Numbers (1–20: 1–5)	Montessori Numbers (1–20: 11–20)	Math Motion Zoom
Learning App 2	Montessori Numbers (Numerals from Quantity: 1–9)	Montessori Numbers (Numerals from Quantity: 10–99)	Place Value Cards
Post App	Montessori Numbers (Quantity: 1–9)	Montessori Numbers (Quantity: 10–99)	Montessori Numbers (Quantity: 100–999)

References

Adesina, A., Stone, R., Batmaz, F., & Jones, I. (2014). Touch arithmetic: A process-based computer-aided assessment approach for capture of problem solving steps in the context of elementary mathematics. *Computers & Education, 78*, 333—343.

Bartoschek, T., Schwering, A., Li, R., & Münzer, S. (2013). Ori-Gami—An App fostering spatial competency development and spatial learning of children. In D. Vandenbroucke, B. Bucher, & J. Crompvoets (Eds.), *Proceedings of the 15th AGILE international conference on geographic information science*. Leuven, Belgium: Springer-Verlag.

Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education, 94*(5), 765—793.

Bertolo, D., Dinet, J., & Vivian, R. (2014). Reducing cognitive workload during 3D geometry problem solving with an app on iPad. In *Presented at the science and information conference, London* (pp. 896—900). IEEE.

Charman, D. (1999). Issues and impacts of using computer-based assessments (CBAs) for formative assessment. In S. Brown, J. Bull, & P. Race (Eds.), *Computer-assisted assessment in higher education* (pp. 85—94).

Chen, L. L. (2011). Enhancing special needs student's learning with iPad. In *In world conference on e-learning in corporate, government, healthcare, and higher education, 2011* pp. 2324—2330).

Chen, L. L. (2012). Integrating iPad in a special education class: A case study. In *In world conference on e-learning in corporate, government, healthcare, and higher education, 2012* pp. 530—534).

Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning, 6*, 81—89.

Clements, D. H., & Sarama, J. (2010). *Learning trajectories in early mathematics — Sequences of acquisition and teaching* (pp. 1—6). Encyclopedia of Early Childhood Development: Numeracy.

Cohen, B. H. (2013). *Explaining psychological statistics* (4th ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Corbin, J., & Strauss, A. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (4th ed.). Los Angeles: SAGE.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.

DeCuir-Gunby, J. T., Marshall, P. L., & McCulloch, A. W. (2012). Using mixed methods to analyze video data: A mathematics teacher professional development example. *Journal of Mixed Methods Research, 6*(3), 199—216.

Duncan, R. G., Rogat, A. D., & Yarden, J. (2009). A learning progression for deepening students' understandings of modern genetics across the 5th—10th grades. *Journal of Research in Science Teaching, 46*(6), 655—674.

Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology, 18*(1), 7—22.

Dweck, C. (2006). *Mindset: The new psychology of success*. Random House.

Falloon, G. (2013). Young students using iPads: App design and content influences on their learning pathways. *Computers & Education, 68*, 505—521.

Ginsburg, H. P., Jamalian, A., & Creighan, S. (2013). Cognitive guidelines for the design and evaluation of early mathematics software: The example of MathemAntics. In L. D. English, & J. T. Mulligan (Eds.), *Reconceptualizing early mathematics learning* (pp. 83—120). Netherlands: Springer.

Goldin, G. A. (2003). Representation in school mathematics: A unifying research perspective. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 275—285). Reston, VA: NCTM.

Goldin, G. A., & Kaput, J. M. (1996). A joint perspective on the idea of representation in learning and doing mathematics. In L. P. Steffe, P. Nesher, P. Cobb, G. A. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 397—430). Hillsdale, NJ: Erlbaum.

Haydon, T., Hawkins, R., Denune, H., Kimener, L., McCoy, D., & Basham, J. (2012). A comparison of iPads and worksheets on math skills of high school students with emotional disturbance. *Behavioral Disorders, 37*(4), 232—243.

Hepplestone, S., Holden, G., Irwin, B., Parkin, H. J., & Thorpe, L. (2011). Using technology to encourage student engagement with feedback: A literature review. *Research in Learning Technology, 19*(2).

Kapur, M. (2014). Productive failure in learning math. *Cognitive Science, 38*(5), 1008—1022.

Kilic, A. (2013). *An investigation of the use of the iPad and textbooks on the achievement of students with special needs in algebra* (Master's Thesis). New Jersey: Rowan University.

Ladel, S., & Kortenkamp, U. (2013). An activity-theoretic approach to multi-touch tools in early maths learning. *The International Journal for Technology in Mathematics Education, 20*(1).

Lee, H. S., & Liu, O. L. (2010). Assessing learning progression of energy concepts across middle school grades: The knowledge integration perspective. *Science Education, 94*(4), 665—688.

Manches, A., & O'Malley, C. (2012). Tangibles for learning: A representational analysis of physical manipulation. *Personal and Ubiquitous Computing, 16*, 405—419. http://dx.doi.org/10.1007/s00779-011-0406-0.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation* (3rd ed.). San Francisco, CA: Jossey-Bass.

Moyer, P. S., Bolyard, J. J., & Spikell, M. A. (2002). What are virtual manipulatives? *Teaching Children Mathematics, 8*(6), 372—377.

Moyer-Packenham, P. S., Anderson, K. L., Shumway, J. F., Tucker, S., Westenskow, A., Boyer- Thurgood, J., et al. (2014, January). Developing research tools for young children's interactions with mathematics apps on the iPad. In *Proceedings of the 12th Annual Hawaii International Conference on Education (HICE), (pp. 1685—1694), Honolulu, Hawaii*. ISSN# 1541-5880.

Moyer-Packenham, P. S., & Bolyard, J. J. (2016). Revisiting the definition of a virtual manipulative. In P. Moyer-Packenham (Ed.), *International perspectives on teaching and learning mathematics with virtual manipulatives* (pp. 5—16). New York: Springer.

Moyer-Packenham, P. S., Bullock, E. P., Shumway, J. F., Tucker, S. I., Watts, C., Westenskow, A., … Jordan, K. (2015a). The role of affordances in children's learning performance and efficiency when using virtual manipulative mathematics touch-screen apps. *Mathematics Education Research Journal*, 1—27. http://dx.doi.org/10.1007/s13394-015-0161-z.

Moyer-Packenham, P. S., Shumway, J. F., Bullock, E., Anderson-Pence, K., Westenskow, A., Boyer-Thurgood, J. M., … Jordan, K. (2015b). Young children's learning performance and efficiency when using virtual manipulative mathematics iPad apps. *Journal of Computers in Mathematics and Science Teaching, 34*(1), 41—69.

Moyer-Packenham, P. S., & Suh, J. M. (2012). Learning mathematics with technology: The influence of virtual manipulatives on different achievement groups. *Journal of Computers in Mathematics and Science Teaching, 31*(1), 39—59.

Moyer-Packenham, P. S., & Westenskow, A. (2013). Effects of virtual manipulatives on student achievement and mathematics learning. *International Journal of Virtual and Personal Learning Environments, 4*(3), 35—50.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Piaget, J. (1946/1970). *The child's conception of movement and speed*. New York: Basic Books (Original work published 1946).

Pirie, S., & Kieren, T. (1994). Growth in mathematical understanding: How can we characterize it and how can we represent it? *Educational Studies in Mathematics, 26*(2), 165—190.

Plummer, J. D., & Krajcik, J. (2010). Building a learning progression for celestial motion: Elementary levels from an earth-based perspective. *Journal of Research in Science Teaching, 47*(7), 768—787.

Reid, D., & Ostashewski, N. (2011). iPads in the classroom—new technologies, old issues: Are they worth the effort? in world conference on educational multimedia. *Hypermedia and Telecommunications, 2011*, 1689—1694.

Revelle, G. L., & Strommen, E. (1990). The effects of practice and input device used on young children's computer control. *Journal of Computing in Childhood Education, 2*(1), 33—41.

Riconscente, M. M. (2012). *Mobile learning game improves 5th graders' fractions knowledge and attitudes* (p. 46). Los Angeles, California: GameDesk Institute.

Riconscente, M. M. (2013). Results from a controlled study of the iPad fractions game Motion Math. *Games and Culture, 8*(4), 186—214.

Roschelle, J. (2000). Choosing and using video equipment for data collection. In A. E. Kelly, & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 709—729). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). Thousand Oaks, CA: Sage.

Sarama, J., Clements, D. H., Barrett, J., Van Dine, D. W., & McDonel, J. S. (2011). Evaluation of a learning trajectory for length in the early years. *ZDM Mathematics Education, 43*, 667—680.

Schubert, C. (2009). Video analysis of practice and the practice of video analysis. In H. Knoblauch, B. Schnettler, J. Raab, & H. G. Soeffner (Eds.), *Video analysis methodology and methods: Qualitative audiovisual data analysis and sociology* (pp. 115—126). New York, NY: Peter Lang.

Segal, A. (2011). *Do gestural interfaces promote thinking? Embodied interaction: Congruent gestures and direct touch promote performance in math* (Doctoral Dissertation). New York: Columbia University.

Smith, C. L., Wiser, M., Anderson, C. W., & Krajcik, J. (2006). Implications of research on children's learning for standards and assessment: A proposed learning progression for matter and the atomic-molecular theory. *Measurement, 4*(1/2), 1—98.

Spencer, P. (2013). iPads: Improving numeracy learning in the early years. In V. Steinle, L. Ball, & C. Bardini (Eds.), *Mathematics education: Yesterday, today, and tomorrow* (pp. 610—617). Melbourne, Australia: MERGA.

Stebbins, R. A. (2001). *Exploratory research in the social sciences*. Thousand Oaks, CA: Sage.

Stevens, S. Y., Delgado, C., & Krajcik, J. S. (2010). Developing a hypothetical multi-dimensional learning progression for the nature of matter. *Journal of Research in Science Teaching, 47*(6), 687—715.

Stroup, W. M., & Wilensky, U. (2000). Assessing learning as emergent phenomena: Moving constructivist statistics beyond the bell curve. In A. E. Kelly, & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 877—911). Mahwah, NJ: Erlbaum.

Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics, 12*, 151—169.

Tashakkori, A., & Teddlie, C. (2010). *Sage handbook of mixed methods in social & behavioral research* (2nd ed.). Thousand Oaks, CA: Sage.

Warshauer, H. K. (2015). Strategies to support productive struggle. *Mathematics Teaching in the Middle School, 20*(7), 390–393.

Winick, D. M., Avallone, A. P., Smith, C. E., & Crovo, M. (2008). *Science framework for the 2009 national assessment of educational progress*. Retrieved from: www.nagb.org/publications/frameworks/science/2009-science-framework.html.

Yin, R. K. (2009). *Case study research design and methods*. Thousand Oaks, CA: Sage.