

UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

Connecting Learning and Developmental Sciences to Educational Preschool Apps: Analyzing App Design Features and Testing Their Effectiveness

Permalink

<https://escholarship.org/uc/item/4rd5s8r1>

Author

Callaghan, Melissa N

Publication Date

2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Connecting Learning and Developmental Sciences to Educational Preschool Apps:

Analyzing App Design Features and Testing Their Effectiveness

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Education

by

Melissa N. Callaghan

Dissertation Committee:
Associate Professor Stephanie M. Reich, Chair
Chancellor's Professor Carol Connor
Associate Professor Gillian Hayes

2018

DEDICATION

To my parents, for giving me everything I needed to get to graduate school.
And, to my husband, for giving me everything I needed to get through it.

TABLE OF CONTENTS

	Page
LIST OF FIGURES	iv
LIST OF TABLES	v
ACKNOWLEDGEMENTS	vi
CURRICULUM VITAE	vii
ABSTRACT:	
Connecting Learning and Developmental Sciences to Educational Preschool Apps: Analyzing App Design Features and Testing Their Effectiveness	x
CHAPTER 1: Introduction	
Applying a Developmental Lens to Educational Game Design for Preschoolers	1
CHAPTER 2: Study 1	
Are Educational Preschool Apps Designed to Teach? An Analysis of the App Market	7
CHAPTER 3: Study 2	
Which App Feedback and Leveling Designs Best Scaffold Preschool Learning?	28
CHAPTER 4: Conclusion	
How Can Digital Apps Support Early Childhood Learning?	48
REFERENCES	58
APPENDIX A: Study 1 Coding Rubric	67
APPENDIX B: Tables of Study 1 Results	82
APPENDIX C: Participants' Background Questionnaire	86
APPENDIX D: Study 2 Game Details	88
APPENDIX E: Study 2 Results Tables	90

LIST OF FIGURES

	Page
Figure 3.1 Froodle Introduction Scene	37
Figure 3.2 Sorting Task Example	37

LIST OF TABLES

		Page
Table 2.1	Sample Feedback Codes	15
Table 2.2	Sample Leveling Codes	17
Table 3.1	Participant Demographics	33
Table 3.2	Group Counterbalancing	35
Table 3.3	Examples of Feedback Types	38
Table 3.4	Scaffolded Leveling Design	39
Table 3.5	Means and Standard Deviations of Participants' Errors in Feedback Tasks	41
Table 3.6	Means and Standard Deviations of Participants' Response Times for Feedback Tasks	42
Table 3.7	Means and Standard Deviations of Participants' Errors in Leveling Tasks	43
Table 3.8	Means and Standard Deviations of Participants' Response Times for Leveling Tasks	44

ACKNOWLEDGEMENTS

I would like to express my sincerest appreciation to all those who have helped make this dissertation research and doctorate degree possible.

A special thank you goes to my committee chair, Dr. Stephanie M. Reich. Her unwavering support to explore a wide-variety of academic endeavors, persist through what feels impossible, maintain a work-life balance, and remain engaged in community outreach, has inspired me to always keep “the bigger picture” in mind. I truly could not have asked for a better graduate school advisor.

I would also like to thank my dissertation committee members, Dr. Carol Connor and Dr. Gillian Hayes, who have helped elevate the quality of this dissertation research to encompass foundational work from learning and computer scientists, allowing me to produce research that could be useful for developers, educators, and consumers of educational technologies.

A warm thank you also goes to Dr. Bill Head, my friend and guiding voice since I started my research career as an undergraduate McNair Scholar. He saw a potential in me that I never imagined existed and he continues to be a powerful role model, teaching me to always push life’s limits.

Lastly, I would like to express my appreciation for the financial support of the National Science Foundation Graduate Research Fellowship (Grant DGE-1321846, 2014) and the Society for Research in Child Development Student and Early Career Council’s Dissertation Research Funding Award. Any opinions, findings, and conclusions or recommendations expressed in this material are those of myself, and do not necessarily reflect the views of the National Science Foundation or the Society for Research in Child Development.

CURRICULUM VITAE

Melissa N. Callaghan
(surname: Powell)

Education

- Doctorate of Philosophy** *September 2013 – June 2018*
University of California, Irvine – School of Education
- Master of Arts Degree** *September 2013 – December 2016*
University of California, Irvine (UCI) – School of Education
- Bachelor of Arts Degree** *August 2008 - May 2013*
California State University, Monterey Bay (CSUMB) – Psychology Department

Honors and Awards

- UCI School of Education Michael E. Martinez Prize, Honorable Mention (2018)**
Recognition annually provided to a doctoral student for superior scholarship, intellectual curiosity, and high levels of professionalism, service, and collaboration.
- SRCD Student and Early Career Council's Dissertation Award Recipient (2017-2018)**
An international dissertation funding award annually given to 10 doctoral students conducting quality research contributing to the field of child development.
- National Science Foundation Graduate Research Fellowship Recipient (2015-2018)**
A three-year \$46,000 doctoral fellowship for conducting innovative research in science, technology, engineering, and math disciplines.
- UCI Multi-Disciplinary Projects Fellow (2016-2017)**
Selected to be funded for a collaborative research project with graduate and undergraduate students from multiple disciplines across UCI campus.
- UCI Campus-Wide Graduate Opportunities Fellowship Recipient (2013-2018)**
A multi-year fellowship that funds Ph.D. students who bring diversity within the university and show promise of completing the program.
- McNair Scholars Program (2011-2013)**
An undergraduate pre-doctoral program which funds first-generation, underrepresented students' research and professional development.

Publications

- Callaghan, M. N.**, Long, J., vanEs, E., Reich, S. M., & Rutherford, T. (2017). How Teachers Integrate a Math Computer Game: Professional Development Use, Teaching Practices, and Effects on Student Achievement. *Journal of Computer Assisted Learning*, 34(1), 10-19. DOI: 10.1111/jcal.12209
- Callaghan, M. N.** & Reich, S. M. (revisions under review). Are Educational Preschool Apps Designed to Teach? An Analysis of the App Market.
- Nikoyalev, M., Reich, S. M., Musket, T., Tadjbakhsh, N., **Callaghan, M. N.** (revisions under review). The use of feedback in online “educational” games for preschoolers.
- Callaghan, M. N.** & Reich, S. M. (under review). Applying a Developmental Lens to Educational App Design for Preschoolers.

Reyes, M., **Callaghan, M. N.**, & McDaniel, A. (under review). High School Students' STEM Occupational Aspirations Reflect Gender Inequality in STEM Occupations.

Research Presentations

Oral Presentations

Callaghan, M. N., Lee, J., Sidhu, A., & Reich, S. M. (2018). Testing the effects of scaffolded app feedback and challenge on preschool learning. American Psychological Association Conference, Washington, D.C.

Yau, J., **Callaghan, M. N.**, Reich, S. M., Wang, Y., Mark, G. (2018). Using Ecological Momentary Assessment and logged data to explore emerging adults' mobile use, mood, and stress. Society for Research on Adolescence Biennial Meeting, Minneapolis Hotel.

Callaghan, M. N., Reich, S. M., & Nguyen, P. (2017). Are educational preschool apps actually educational? A content analysis of math apps on the market. Jean Piaget Society's Technologies and Human Development Conference, San Francisco, CA

Callaghan, M. N., Yau, J., Reich, S. M., Wang, Y., & Mark, G. (2017) The impact of social gaming on mood and stress. Society of Community Action Research Biennial Conference, University of Ottawa, Ottawa, Ontario, Canada.

Reyes, M., **Powell, M.**, & McDaniel, A. (2016). Tracking Students' Pathways to Male- and Female- Dominated STEM Careers: The Influence of Career Expectations, Math Course-Taking and Self-Efficacy. Paper presented at the American Educational Research Association Annual Meeting, Washington, D.C.

Powell, M. & Firek, K. (2014). Family use of sarcasm and teasing and its relation to bullying experiences. Social Science Research & Instructional Council Student Research Conference, California State University, Fullerton, CA

Powell, M. (2013). Examining Girls Inc. College Bound and SMART curriculum. University of California Irvine, Summer Research Program Symposium, Irvine, CA

Powell, M., Haimovitz, K., & Walton, G. (2013). Private speech: A strategy to delay gratification. California State University, Monterey Bay, Senior Capstone, Monterey, CA

Powell, M., Haimovitz, K., & Walton, G. (2012). Willpower: Using private speech to delay gratification. Leadership Alliance National Symposium, Convention Center, CT

Powell, M., Park, H., Manago, A., Greenfield, P., & Dyer-Seymour, J. (2012). Familistic and individualistic values within families of emerging adults across cultures. CSU Student Research Competition, California State University, Long Beach, CA

Poster Presentations

Powell, M. & Reich, S. M. (2016). Applying a developmental lens to preschool game design. Society for Research on Child Development's Technology and Media in Children's Development Conference, University of California, Irvine.

Powell, M., Reich, S. M., Burciaga, E. (2015). Fostering classroom community through student disclosures in college readiness programs. Society for Community Research and Action Conference, University of Massachusetts Lowell, MA

Powell, M., Reich, S. M., Burciaga, E. (2014). 'I'm freaking out!' Student disclosures in college readiness programs. SoE Poster Symposium, University of California, Irvine, CA

Powell, M., Park, H. Manago, A., Greenfield, P. (2012). Familistic and individualistic values within families of emerging adults. Western Psychological Association Conference (WPA), Hyatt Regency San Francisco Airport, Burlingame, CA

Powell, M., Park, H. Manago, A., Greenfield, P. (2012). Familistic and individualistic values within families of emerging adults. Graduate Division Summer Program for Undergraduate Research, University of California, Los Angeles, CA

LEADERSHIP & SERVICE

Peer Mentor for UCI's Campus Wide Summer DECADE program
University of California, Irvine (Summer 2016; Summer 2017)

Writing Mentor for UCI's Campus Wide Summer DECADE program
University of California, Irvine (Summer 2016; Summer 2017)

Co-Leader of the UCI School of Education DECADE Council
University of California, Irvine (Fall 2015-Spring 2017)

Steering Committee Member for UCI School of Education
University of California, Irvine (Fall 2016-Spring 2017)

ADSE Committee Member for UCI School of Education
University of California, Irvine (Fall 2016-Spring 2017)

Peer Mentor for UCI School of Education DECADE
University of California, Irvine (Winter 2015-Spring 2017)

SRCD Conference Program Planning Committee Member
Society for Research in Child Development's Technology and Media in Child Development conference (November 2015-October 2016).

Admissions Committee Member for UCI School of Education
University of California, Irvine (Winter 2015, Spring 2016)

ABSTRACT OF THE DISSERTATION

Connecting Learning and Developmental Sciences to Educational Preschool Apps:
Analyzing App Design Features and Testing Their Effectiveness

By

Melissa N. Callaghan

Doctor of Philosophy in Education

University of California, Irvine, 2018

Associate Professor Stephanie M. Reich, Chair

To date, applying what is known about child development to educational app design has been minimally studied and rarely tested. Using developmental and learning sciences research (e.g., Hirsh-Pasek et al., 2015), this dissertation identifies whether evidence-based teaching strategies are used in the design of existing educational apps and experimentally tests how some tactics are more effective than others. Study 1 aims to understand how popular preschool literacy and math apps are designed and whether they connect with evidence-based teaching practices. This content analysis of 171 top literacy and math premium and free apps on Apple and Android app markets uses a defined spectrum of effective teaching tactics for young children identified in developmental and learning sciences (e.g., feedback, increasing complexity, guided play), to identify whether these tactics have been applied to current apps' mechanics. Findings show there is a wide variety of features used across preschool apps, but few provide developmentally appropriate guidance for young learners, such as scaffolded feedback and leveling. Study 2 explores the direct effects of these features with 240 preschoolers, to experimentally test: a) which types of audible app feedback best support preschool learning, and b) whether

preschoolers using an educational app with scaffolded challenge (i.e., leveling up) learn better than preschoolers using the same app with random ordered challenge. An ANCOVA and post-hoc Tukey test show that, of the participants who played the feedback tasks first (n=120), those with scaffolded feedback (i.e., explanations for correct and incorrect answers) were more accurate than participants with sound responses (i.e., bell, gong). However, this advantage disappeared when children played the leveling game first. Participants with encouraging (i.e., “Great job!” or “Try again!”) feedback took longer to complete tasks, compared to the sounds feedback group. Moreover, leveling results showed that gradually increasing challenge helped preschoolers make fewer incorrect responses and complete tasks faster than children playing the same game with random order leveling. Ultimately, improving the design of educational preschool apps through empirical research, like this, may be what our increasingly digital age needs to provide young children with high quality digital learning tools.

CHAPTER 1: INTRODUCTION

Applying a Developmental Lens to Educational Game Design for Preschoolers

Prior to formal schooling, young children often begin learning through mobile devices in their homes (Plowman, McPake, & Stephen, 2010), which may explain why 58% of educational games (apps) for mobile devices, such as tablets and smartphones, are made for preschool-aged users (Highfield & Goodwin, 2013; Shuler, 2012). These digital learning tools will undoubtedly continue to grow in popularity (Plowman et al., 2012) due to society's consistent advancement in and embracing of digital technologies (Molnar & Cavanagh, 2013). Yet, if educational apps for young children are not designed with an understanding of how children develop and learn (Hirsh-Pasek et al, 2015; Reich & Black, 2012), educators and caregivers could waste valuable time, money, and resources on products that do not actually teach their children (Cuban, 2001; Rutherford et al., 2015). This dissertation connects extant learning and developmental sciences research to app design, illuminating how educational apps could be designed to support learners aged 3 to 5 years.

Though apps could be useful learning tools for players of all ages, developing educational apps for preschool children can be particularly challenging, since young children process information differently than older individuals (e.g., shorter attention-spans, still developing reading skills, etc.; Hirsh-Pasek et al, 2015). However, since preschool learning is crucial for later educational success (Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002), and research shows that interweaving play with teaching is desirable among preschool-aged learners (Gmitrova, Podhajecka, & Gmitrova, 2009), apps could be a very useful platform for teaching this young age group. Yet, to be such, designers must consider how preschool-aged users typically learn.

Though there are often individual differences in how children learn, research has helped identify teaching strategies that are typically effective for preschoolers in face-to-face settings and when using digital technology. Importantly, researchers have previously explored how digital media have the capacity to connect to key factors of child learning (Chiasson & Gutwin, 2005; Cooper, 2005; Hirsh-Pasek et al., 2015; Kankaanranta et al., 2017; Lieberman, Bates, & So, 2009). This dissertation extends such prior work by merging these key teaching strategies for preschool-aged audiences and applying them specifically to educational preschool app design.

This chapter will briefly discuss how the two studies within this dissertation will address gaps in current published research. Chapter 2 will then identify whether developmentally appropriate, evidence-based practices for teaching preschoolers are actually utilized by popular (i.e., top rated) math and literacy preschool apps on Apple and Android app markets. Following that, Chapter 3 will describe an experiment designed to test the effectiveness of two core design features utilized by preschool apps on the market – feedback and leveling. While the introductions for Chapters 2 and 3 may seem redundant because they are written in manuscript format, each chapter aims to use differing methods to capture and emphasize ways in which educational apps can be designed to support preschool learning. Lastly, Chapter 4 will bridge each of these dissertation chapters into a cohesive plan for how current and future educational preschool apps can be designed to support the young, unique minds of preschool-aged learners.

Why Educational Preschool Apps are Important to Study

Both studies of this dissertation utilize past developmental and learning sciences research as a foundation for understanding and designing educational preschool apps. By empirically measuring the features used across today's preschool app market and experimentally testing the direct effects of certain preschool app features, the two studies of this dissertation aim to fill in

some of the current gaps in published literature. The following subsections will identify these gaps and explain how this dissertation will address them.

Study 1: Understanding How Popular Preschool Apps Are Designed

Previous Analyses of Apps on the Market. To date, there is limited knowledge of whether digital educational apps are designed to meet the unique needs and learning abilities of preschool-aged learners. When examining 200 top-sold Apple apps on the market, Shuler (2012) found that 100% of the apps which claimed to promote learning outcomes were advertised for children ages 5 and under. Yet, with minimal research examining these apps' educational design features, there is little understanding of what preschool-aged learners are playing.

A few researchers have conducted content analyses of children's educational apps, which provide first-step insights to how current apps are designed to teach young learners. When looking at the educational features of Apple math apps for all ages, Highfield and Goodwin (2013) found that many apps used "drill and practice" activities, while some used open-ended tasks; fostering creative learning experiences, without assessing right or wrong answers. When looking specifically at educational apps for preschool-aged learners, two studies (Blair, 2013; Vaala, Ly, & Levine, 2015) show that math and literacy apps either provided no feedback or required players to answer until correct. Few educational preschool apps described reasons behind correct answers. This research could be further extended by considering more educational design features known to be effective in non-digital and digital studies done in the educational and developmental sciences.

Aims of Study 1. Study 1 aims to identify whether prevalent design features of educational preschool apps align with empirical-based teaching practices. Expanding on the work of previous researchers, Study 1 explores the clarity of task prompts, how different types of

feedback (audible, visual, and verbal) are used to promote children's understanding of the educational content, whether levels of challenge are designed to adapt to user-input, and what types of reward systems are used to motivate children to want to do well and continue learning. By measuring the variety of design features utilized by educational preschool apps from the beginning to end of each game, Study 1 provides a holistic picture of whether these apps are designed to support preschool-age learning.

Special Note for Study 1. Chapter 2's manuscript currently has revisions under review for the Journal of Learning, Media and Technology.

Study 2: Testing the Effects of Preschool App Feedback and Leveling Designs

App Design to Promote Preschool Learning. In addition to documenting the current educational app design features in top rated apps, the next crucial step is experimentally testing apps to determine which design features are most effective for teaching preschool-age learners. While every feature of educational apps may be important for promoting preschool learning, feedback can particularly be used to help learners better understand the game's educational content and leveling can guide learners toward understanding increasingly more challenging content. Yet, few researchers have examined which designs of app feedback and leveling are most effective for preschool-aged learners. One pilot study (Blair, 2013) examined preschoolers' strategies of play based on the types of app feedback they received. Findings showed that children explored more using trial-and-error when they were forced to keep answering until correct, whereas children took more time for deliberate decisions when apps corrected wrong answers or further described their answers. Still, little is known about which types of feedback better support preschoolers' learning from educational apps. Additionally, though research suggests guiding young learners toward understanding complex content is more useful than

having preschoolers learn solely through trial and error (Hsin & Wu, 2011; Henderson, Many, Wellborn, & Ward, 2002), minimal research has been done to determine whether this is also true for digital app learning. Since the educational impact of well-designed feedback and leveling has the potential to guide preschoolers toward understanding new and complex content, experimental tests can help inform the design of educational preschool apps to make these digital tools truly effective.

Aims of Study 2. The goal of Study 2 is to determine which feedback and leveling designs best support preschoolers in learning from digital apps. While researchers have explored how different types of app feedback impact children’s patterns of play, Study 2 focuses on whether there are certain types of audible feedback that help children make fewer incorrect responses. This is done by exploring (1) whether sound effects (i.e., bells/horns) that solely indicate if children answered tasks correctly or incorrectly guides preschool learners to making minimal mistakes, (2) if brief verbal responses that praise and encourage players (i.e., “Great job!” and “Try again!”) help preschoolers make fewer mistakes, or (3) whether it is the feedback that concretely explains the reasons for answers being correct or incorrect that helps young learners make fewest mistakes. Likewise, Study 2 also explores whether certain types of leveling better support preschoolers in making fewer mistakes, by comparing the effects of an app with no strategy in leveling design (randomly ordered challenge) to the same app with levels that gradually increase in challenge. These findings can provide guidelines for how current and future educational preschool apps may be designed to effectively support young learners.

Conclusion

Designing educational apps that capitalize on the developmental needs of preschoolers has the potential to increase the effectiveness of these digital learning tools for 3- to 5-year-old

children who may use them in schools, homes, and other out-of-school environments.

Educational apps have the potential to enhance preschool education through effective and engaging play. Since there is already a high percentage of educational preschool apps available on the app market today (Highfield & Goodwin, 2013; Shuler, 2012) and mobile devices are in almost every young child's home (Pew Research Center, 2017), a vital first step is to determine whether today's apps are designed with a developmental lens to support preschool-aged learners. Therefore, the following chapter discusses a content analysis study examining popular preschool math and literacy apps on the market.

CHAPTER 2: STUDY 1

Are Educational Preschool Apps Designed to Teach? An Analysis of the App Market

Young children's exposure to digital technology is increasing, and there has been a growing interest in utilizing this technology to create innovative educational tools. This has become especially prevalent across mobile application (app) markets, with over half of educational apps advertised for preschool-aged learners (Highfield & Goodwin, 2013; Shuler, 2012). This increase in popularity is likely due to the intrigue of learning from a playful game and the personalization features that digital apps can be programmed to offer. For example, educational apps can adapt content and challenge to align with the current knowledge and abilities of individual players. Content can grow more challenging as learners do well on an app or decrease in challenge when learners appear to consistently struggle (e.g., answer incorrectly multiple times in a row), using preprogrammed algorithms (e.g., Thorell et al, 2009). Mobile devices also offer interactive features that may be used to support deeper learning for young children. For example, device movements may be used to teach physical concepts (e.g., shaking the device to teach force, tipping the device to simulate the pouring of liquid measurements, etc.) and device cameras could be used to integrate pictures of "real world" objects to draw connections between in-game content with learners' out-of-game worlds. Moreover, since 84% of homes have at least one mobile device available (Pew Research Center, 2017) and apps are often sold at a relatively low cost (Shuler, 2012), there are increasing opportunities for learners from diverse socioeconomic backgrounds to access them. Regardless of the growing capabilities and accessibility of educational apps, these digital tools for young learners have been minimally studied.

Years of developmental research have helped to identify key elements for teaching

preschool-aged learners, but little research has examined whether such elements are present within educational apps on the market. This is especially concerning since preschoolers' learning is foundational for later academic success (Campbell et al, 2002), and young children process information so differently from older individuals (e.g., shorter attention spans, more limitations in visual processing, still developing working memory; Hourcade, 2007) and are still developing their fine-motor skills needed for accurate touchscreen interactions (Vatavu, Cramariuc, & Schipor, 2015). If current educational apps are not designed to align with the developmental abilities of preschool-aged learners, utilizing app technology to reach and teach a wide range of learners may not actually be educational. Though young children are continuously surrounded by such technology (Plowman et al, 2012), there is little guarantee of whether the educational apps that preschoolers play are tailored to meet their developmental needs and abilities.

This study examines the most popular educational preschool apps on the market and whether their features align with what researchers have identified as typical capabilities of preschool-aged users and best practices for teaching this age group. More specifically, this study seeks to answer: How are popular preschool math and literacy apps designed? And, do their features connect with evidence-based teaching practices? With limited knowledge of how such apps are designed to teach, this study sheds light onto the teaching tactics commonly and rarely used across young children's educational apps, ultimately identifying design features that are and are not used to improve preschoolers' educational experiences.

Key Elements for Teaching Preschoolers with Apps

To assess the developmental appropriateness of math and literacy apps currently on app markets, a framework was developed based on extant literature examining how preschool-aged children develop and learn. For decades, researchers have been examining teaching tactics

that could be effective when working with preschool-aged children (Hirsh-Pasek et al, 2015; Hourcade, 2007). Each of the following subsections identifies key design elements of educational preschool apps that were coded in this current study.

Clarity and simplicity of goals. Preschool-aged learners are just beginning to build their cognitive processing skills (Garon, Bryson, & Smith, 2008), so educational apps with clear and simple goals may be more effective in supporting young children's cognitive development. At the start of play, educational apps can provide clear prompts for users to understanding what the task is and how they can complete it. Research has found that young children learn best with clear instructions and modeling that allows them to draw connections to existing beliefs (Kloos & Orden, 2005). For example, Kloos and Orden (2005) found that preschool-aged children's original beliefs about mass and volume could be corrected after seeing a demonstration of circumstances that caused boats to sink. This modeling method can also be utilized by educational apps to help users understand why certain behaviors yield certain outcomes, providing a deeper causal explanation for the educational task(s). Apps can also repeat and/or rephrase instructions after pauses in user play, to help shift and maintain young children's attention (Kannass & Colombo, 2007) on the desired goals of the app. Therefore, this current study's coding framework measures whether apps use clear prompts and/or character demonstrations to model correct and incorrect ways of achieving the goal(s) of the task, and in-play guidance via modeling and repeated/rephrased instructions. The cognitive support preschool-aged learners could receive from well-designed prompting may be further enhanced if it is followed by well-designed feedback.

Quality of feedback and rewards. Though researchers are still striving to identify what "high quality app feedback" for young children entails, developmental and learning

sciences research suggests that certain types of feedback are more likely to be useful for preschoolers than others. For example, since preschool-aged children are just beginning to develop their reading skills (Lonigan, Burgess, & Anthony, 2000), providing feedback in the form of text would likely not help this age group, as they would likely struggle to comprehend it. While audible feedback (e.g., bells and horns) may help a child to know whether an answer is correct or incorrect (Blair, 2013), sound effect feedback alone lacks further indication of *why* something is correct or incorrect. Apps have the capacity to visually demonstrate and verbally describe why preschoolers' actions yielded certain outcomes (Gavriushenko, Karilainen, & Kankaanranta, 2015). This explanatory form of guidance could be seen as scaffolding feedback. Scaffolding is often defined as guidance, which carries a learner from a basic understanding of content to gradually understanding more complex and difficult content (Hsin & Wu, 2011; Obikwelu & Sin, 2012; & Vygotsky, 1978). It is important to know which types of feedback are prevalent across popular math and literacy preschool apps on the mobile app market, gauging whether apps use explanatory feedback for scaffolding young learners to understand the educational content being taught.

Rewards provided at the completion of tasks can also be designed to guide learning. Badges, stickers, and high scores are often used by apps as incentives to succeed in completing educational tasks and progressing further into the game (Zichermann & Cunningham, 2011); however, research shows that these forms of gamification actually undermine learners' motivation to understand and progress through the app's educational content (Botha, Herselman, & Ford, 2014; Deci, Koestner, & Ryan, 2001). Instead, researchers (e.g., Zichermann & Cunningham, 2011) suggest that educational games, such as preschool apps, should utilize "learning" (e.g., unlocking new levels of a game), as the desired reward children should strive

for as they progress to advanced levels.

Structure of challenge. App leveling can be designed strategically to scaffold young children's learning to gradually increase in difficulty as preschoolers progressively understand the material, as well as decrease in difficulty when preschoolers appear to struggle. Number knower levels, cardinality, letter recognition, and reading comprehension vary greatly during the preschool years (Lonigan, Burgess, & Anthony, 2000; Sarnecka, 2015), making a 3-year-old learner potentially very different from a 5-year-old learner. Utilizing adaptive scaffolded challenge, apps could be purposefully designed to help children build a foundational understanding of basic concepts before engaging with more sophisticated, complex material. In addition to apps' ability to adapt to user input, mobile devices could provide physical interactivity features that also foster learning.

Motion-based interactions. After considering the physical capabilities of young learners, app designers may want to take advantage of the innovative opportunities made possible with mobile devices. As researchers have found in the past, physical manipulatives can assist young children in deeply understanding topics, such as counting (Petersen & McNeil, 2013), but complex touchscreen interactions can be challenging for young learners who are still developing their fine motor skills (Vatavu, Cramariuc, & Schipor, 2015). Ensuring that the interface of educational apps (e.g., touchable object sizes, simplified touchscreen motions at first, etc.) align with the physical capabilities of preschool-aged learners can help in creating a smoother play experience (Romeo et al, 2003). The unique device features of mobile devices can also be used to support conceptual understanding. For instance, accelerometer technology could be used to support learning about volume and measurement, such as tipping of a device to virtually pour liquid from one container to another, with varying times based on amount (Piaget,

1984). To date, it is unclear whether such innovative uses of device manipulations are incorporated into preschool educational app play.

If these physical device opportunities, and the many other design opportunities discussed in this review, are not taken advantage of, apps may not be designed to their fullest educational potential. Therefore, it is important to know how educational preschool apps on the market are currently designed to teach preschool-aged learners.

Educational Apps Reviewed to Date

Over half of educational apps on Apple, Google, and Amazon app markets are advertised for users 5 years and under (Shuler, 2012; Vaala et al, 2015), yet there is little indication of whether the key elements of learning discussed above are present within these preschool apps. The following subsection describes what is known about educational preschool app design.

Analyses of apps' educational content. When examining the educational structure and content of math apps on the Apple market for all ages, Highfield and Goodwin (2013) found that 89% of the apps reviewed (n=47) were designed using “drill and practice” types of activities and the six remaining apps used manipulability and instructive features. Some apps had an open-ended nature that did not assess right and wrong responses, fostering opportunities for children to practice creativity, without directly teaching correct/incorrect ways of manipulating the educational content. Blair (2013) further supported these findings when analyzing specific forms of feedback used across 50 preschool math apps. She found that most of the math apps (87%) either provided no feedback or required players to play the same task until they got it correct. Vaala and colleagues (2015) similarly found that literacy preschool apps only provided feedback 10 to 23% of the time. When looking at the type of feedback provided by apps, Blair (2013) found that very few preschool math apps (less than 7%) provided elaborate feedback on correct

answers for players to understand why they were correct. These studies (Blair, 2013; Highfield & Goodwin, 2013; Shuler, 2012) provide some insights into how mobile apps are designed to teach young learners. Yet, there is still little known about the collection of design features used in educational games, including instructions, feedback, rewards, and challenge that are aimed at teaching preschoolers through mobile devices.

Current Study

This study seeks to identify the common and missing educational design features of math and literacy apps for preschoolers. With little known about the design features of preschool apps marketed as educational or how their design connects to our extant understanding of how 3-5-year-old children learn, this content analysis of the Apple and Android app markets examines: How popular preschool math and literacy apps are designed and how their design features connect with evidence-based teaching practices.

Method

App Selection. Each month, for three months, the top 10 premium (paid) and 10 free apps from Apple and Android (Amazon and Google Play) stores were selected. This was done for math apps (June to August 2016) and literacy apps (April to June 2017). These apps were categorized by the App stores as educational games for players “under the age of 5” and had math or literacy learning outcomes listed in the game description and/or title. Only games that assessed correct and incorrect responses were included, as they had a clear educational goal in which players could theoretically learn and improve performance over time. In total, 171 apps (62 math, 60 literacy, and 49 both math and literacy) were included in the study (note that some apps persisted in the Top 10 for more than one month).

Conventional and Directed Content Coding. Two different approaches to *content analysis* were used - conventional and directed. Though the research team relied on preliminary game play to create various codes for our framework, we also utilized previously published educational and developmental research (e.g., Hirsh-Pasek et al., 2015) to create codes for teaching tactics that have been identified as important for teaching preschoolers and previously established playability heuristics (e.g., Shoukry et al., 2015). These two content analysis approaches were combined to allow preschool games' content to naturally shape how and what we measured, while also applying existing theory as a way of specifically targeting the educational features of games (Hsieh & Shannon, 2005). Coders assessed app responses for when answering correctly and incorrectly, and also sat idle during each task for at least 30 seconds to see if the app provided additional visuals or audible directions when users stop interacting with the app.

Development of Codes. In order to capture the types of educational features commonly used within apps, a coding framework was developed by five researchers, drawing on a coding scheme developed for a content analysis of educational preschool online computer games (Nikolayev et al, under review). The framework was then modified based on preliminary play of preschool games on computers and tablets, and a defined spectrum of learning tools identified in developmental and learning sciences (e.g., feedback, increasing complexity, guided play) as effective for young children's learning (Clements & Sarama, 2002; Hirsh-Pasek et al., 2015; Hourcade, 2007). To also account for features of the apps that may impact children's ease of play and enjoyment of the game, design heuristics that are well established within Computer Sciences and Human Technology Interactions research were used to create additional codes (e.g., Shoukry, Sturm, & Galal-Edeen, 2015).

Coding Framework. Once a basic coding framework was created, a research team of graduate students, undergraduate students, and a professor tested the framework and collaboratively modified codes to better capture the various learning features used within game mechanics. The final version of the coding framework documented concepts, such as character design (e.g., personalized avatar vs. pre-made characters, first person vs. third person view), targeted math (e.g., patterns, addition/subtraction, counting) and literacy subjects (e.g., alphabetical order, trace letters, picture naming), responses to various forms of input (e.g., correct vs. incorrect answers, consistent incorrect answers) or lack of input (e.g., repeating or rephrasing instructions), reward systems (e.g., stickers, leveling up), and more. Some apps included multiple mini-games (e.g., counting game, addition game, and subtraction game), usually separately accessible through a menu. For these, we coded each mini-game as a separate “puzzle”. For apps that only had one game, they were coded as having “1 puzzle.” Codes included both commonly used and scarcely used design features, and each puzzle of each app was separately coded (unless it had more than 10 puzzles, in which case, we only coded the first 10 puzzles of the app). The entire coding framework can be found in Appendix A, and examples of feedback (Table 2.1) and leveling (Table 2.2) codes can be found below. After coding began, if there were game features found that did not fall within an existing code, those unique design features were documented as “additional notes”; though, this rarely occurred.

Table 2.1. Sample Feedback Codes (see Appendix A for full coding rubric)

NON-VERBAL, SOUND FEEDBACK	
When Correct	When Incorrect
No sound effect	No sound effect
Small sound (e.g., bell)	Small sound (e.g., gong)

Ambiguous emotional sound (e.g., sound from a character that is difficult to distinguish whether it is positive or negative, such as a screeching giggle)	Ambiguous emotional sound (e.g., monkey sound that is difficult to distinguish whether it is positive or negative)
Positive emotional sound (e.g., cheers)	Elaborate negative sound (e.g., waa-waa)
Elaborate sound (e.g., music)	
VISUAL FEEDBACK	
When Correct	When Incorrect
Nothing changes	Nothing changes
Demonstrates answer was right (e.g., answer lights up in green)	Non-scaffolding feedback (e.g., a character simply shifts position)
Positive emotion/praise (e.g., confetti and balloons)	Demonstrates answer was wrong (e.g., “X” over selection or highlights in red)
Neutral changes that are not scaffolding (e.g., an animal moves to a different side of the screen for no clear purpose)	Demonstrates correct answer (e.g., correct answer flashes)
General guiding hints (e.g., dotted lines to show distance between objects)	Subtle negative emotional feedback (e.g., sad faces)
Scaffolded visual feedback (e.g., highlights oranges one by one to count the total selected)	Scaffolding via general guiding hints to reach correct answer (e.g., crosses out wrong answers leaving fewer possible choices, etc.)
VERBAL FEEDBACK	
When Correct	When Incorrect
No verbal feedback	No verbal feedback
Positive praise (e.g., “Great job! You did it!?”)	Non-scaffolded verbal feedback (e.g., “Not quite. Try again.”)
Narrates action, stating correct action (e.g., “You selected three oranges! Correct!”)	Gives correct answer (e.g., “There are three oranges.”)
Scaffolded feedback (e.g., “One...two...three! That’s three oranges!”)	Scaffolded feedback (e.g., “Not quite. A circle has no angles. You selected a square, which has 4 angles.”)

Unrelated feedback (e.g., “Square starts with the letter ‘S’!”)	Unrelated feedback (e.g., “Find the square. Square starts with the letter ‘s’”)
	Repeated verbatim instructions (e.g., “How many oranges?”)
	Rephrased instructions (e.g., “Count the oranges.”)
TEXTUAL FEEDBACK	
When Correct	When Incorrect
No textual feedback	No textual feedback
Positive praise (e.g., “Yay! You did it!”)	Non-scaffolded feedback (e.g., “Nope.”)
Scaffolded feedback (e.g., “One...two...three. Correct! There are three trucks!”)	Scaffolded feedback (e.g., “One...two...three...There are only three trucks. We need four.”)
Unrelated textual elaboration (e.g., “Boats have triangles too.”)	Unrelated textual feedback (e.g., “Boats have triangles.”)
	Repeated written instructions
	Rephrased written instructions

Table 2.2. Sample Leveling Codes

CHALLENGE WITHIN EACH PUZZLE
There was no scaffolded challenge in this puzzle. Level of difficulty remained constant.
Challenge decreased within puzzle after failures (e.g., after getting many addition problems with high quantities wrong, puzzle transitioned to smaller numbers in addition problems)
Challenge increased within puzzle after successes (e.g., after getting basic addition problems correct, quantities increased)
Increase in challenge/reward process (e.g., more difficult to collect items/points)
LEVELING ACROSS WHOLE GAME
Not present
Subtle levels of play that increase in challenge (e.g., No indication of level numbers, but

gradually increases in challenge)
Obvious levels of play that increase in challenge (e.g., number of points/puzzles leads to increased level number)
Level of difficulty can be selected (e.g., easy, medium, hard)
SCAFFOLDING ACROSS WHOLE GAME
Level of challenge did not appear to increase or decrease in challenge ever
Challenge increased across puzzles (e.g., first puzzle was counting, second puzzle was addition, etc.)
Challenge decreased across puzzles (e.g., after getting number recognition problems wrong, it transitions to puzzles with number tracing)
Increase in non-math related challenge (e.g., more difficult to collect items/points)

To present results in a clear, cohesive manner, these codes were later placed within categories representing the key educational elements of design previously discussed. These are 1) Simplicity and Clarity of Goals (learning goals and features of enjoyment, types of prompts, modeling goals, in-play guidance), 2) Feedback and Rewards (consistency of guidance, quality of feedback, reward systems), 3) Structure of Challenge (leveling, scaffolded challenge), and 4) Mobile App-Based Interactions (touchscreen interactions, mobile-device capabilities).

Inter-rater reliability. Establishing reliability across the five coders was done over the course of several weeks. Pilot coding was first done together, discussing where certain game features would fall within the coding framework. Then, additional pilot coding was done separately, eventually comparing codes to one another and discussing discrepancies. Finally, by collectively coding one game every two weeks, coders' inter-rater reliability reached a 0.80 to 0.85 level of reliability across categories. Every two weeks, coders reestablished this level of reliability for all coding categories, using one new app for each calibration, and any

discrepancies found were discussed in detail before the researchers returned to coding the remaining apps for the study.

Analysis

Initially, overarching patterns were identified through frequency counts to describe the features commonly and rarely used in apps. Each coded characteristic was dichotomized (0 = absent, 1 = present) to indicate whether the code was or was not present within the app. Frequencies and chi-square statistics were calculated, identifying the prevalence and patterns of co-occurring features (e.g., how frequently do apps use explanatory feedback and is it more commonly used with scaffolded leveling?). Additionally, comparisons between target content (math vs. literacy), app type (free vs. premium) and tablet platform (Apple vs. Android) were calculated. Tables of each results section can be found in Appendix B.

Results

Simplicity of Content and Clarity of Goals

Learning goals and features of enjoyment. Overall, most apps (79.5%) presented clear goals to users, with a single theme across play (87.1%). Of the math apps, the most common learning goals were counting numbers or objects (50.5%) and number recognition (40.5%). The least common math learning goals were even/odd numbers (0.01%) and currency (0.01%). Of the literacy apps, the most common learning goals were identifying letters (33.0%) and phonological awareness (26.6%) (e.g., sounding out letters, adding letters to create a word). Very few apps taught differentiating vowels and consonants (0.2%), story comprehension (0.01%), and completing sentences (0.01%).

Types of prompts. While some apps (18.7%) did not provide instructions, requiring players to learn the goals of the game through trial and error, many (32.8%) provided moderate

instructions (e.g., “How many apples are in the basket?”) and some (25.2%) provided additional elaborate instructions on how to complete tasks (e.g., “How many apples are in the basket? Tap on the correct number”). This was consistent when comparing apps’ target content - moderate instructions were most common for math (30.6%) and literacy (28.4%) apps. These prompts were primarily given verbally (79.0%), while 35.7% of those apps coupled text with the verbal prompts. Some (13.5%) of the apps provided visual cues to demonstrate the task’s instructions (e.g., lighting up the apples while counting out loud), with little difference between math (12.6%) and literacy (11%) apps.

Modeling goals. Very few apps provided modeling to demonstrate how to complete tasks. Only 4.7% of the apps used characters to model how to get to correct answers. There were more apps (17.0%) that demonstrated how to complete tasks (e.g., an arrow or hand showing the dragging motions required) and some (10.5%) provided hints to demonstrate next steps (e.g., “When you’re finished, tap on the green button!”). These patterns did not statistically differ between math and literacy apps [χ^2 (2, $N=171$) = 0.95, $p=0.62$], Apple and Android apps [χ^2 (1, $N=171$) = 1.73, $p=0.19$], or free and premium apps [χ^2 (1, $N=171$) = 3.53, $p=0.06$].

In-play guidance. Since children may not understand the initial instructions provided by apps, it was important to examine whether apps frequently repeated or rephrased instructions if there was ever a 30-second pause in play. Few apps (15.2%) repeated instructions after a pause, and even fewer (6.4%) rephrased instructions. Apple apps (11%) rephrased instructions more than Android apps (1%); [χ^2 (1, $N=171$) = 5.95, $p<0.05$], and the premium apps (23%) repeated instructions more than the free apps (9%); [χ^2 (1, $N=171$) = 7.26, $p<0.01$].

Quality of Feedback and Rewards

Types of feedback. Overall, positive feedback was common, while scaffolding feedback was rare. A sense of positivity was common for correct answer feedback across all four formats coded – sound effect, verbal, visual, and textual feedback. Of the sound-based feedback, 28.7% of apps provided small sound effects as feedback (e.g., bell or horn), with another 28.1% of apps using positive sound effects, such as celebratory chimes or music. A high number of apps (73.7%) also used positive verbal feedback (e.g., “Great job!”). While most (84.2%) apps did not use textual feedback for correct answers, those that did, used it for positive statements (23.4%). When coded, 73.1% of apps visually demonstrated correct answers (e.g., flashes correct answer in green), and 68.4% also use celebratory visual feedback (e.g., confetti and fireworks).

Very few apps provided feedback for incorrect answers that was structured to scaffold understanding of why the action was wrong or how to improve it. A high number of apps (65.5%) used small sound effects (e.g., gong, buzzer) as feedback. Verbally, most (46.8%) apps used non-scaffolded feedback that did not guide users towards understanding why the selected answer was incorrect, or what made a different answer more correct. Rather than instructive feedback, such as “No, that is a circle. It doesn’t have four corners, like a square. Tap on the square!”, apps responded with “No, try again.” Most apps (94.5%) did not use textual feedback and the only apps that did used non-scaffolded textual responses. Many apps used visual feedback, most (57.3%) using visuals to demonstrate an answer was incorrect (e.g., A red “X” across the selected answer) and several (35.7%) used neutral visuals that did not provide any information about how to reach correct answers, nor why a selection was wrong (e.g., a character blinks).

In comparing the use of scaffolded feedback between math and literacy games, we found that oral and visual scaffolded feedback were relatively equal between apps [Oral: χ^2 (1, $N=171$)

= 0.28, $p=0.6$; Visual: $\chi^2 (1, N=171) = 0.09, p=0.77$]. Few math (17.7%) and literacy (15.0%) apps used oral scaffolded feedback. About half of math (54.8%) and literacy (55.0%) apps utilized visual scaffolded feedback (e.g., items light up in the order they are counted). Scaffolded feedback through text was very uncommon (only 1 math app).

Further, there was no different in the use of scaffolded feedback between free and premium apps (34.0% of free apps, 39.5% of premium apps) [$\chi^2 (1, N=171) = 0.44, p=0.51$]. Scaffolded feedback was used similarly by 32.9% of Android apps and 40% of Apple apps [$\chi^2 (1, N=171) = 0.67, p=0.43$].

Consistency of guidance. Though modeling was rare, it was possible that apps, which used pre-play modeling, also guided users through scaffolded feedback. When examining if apps were consistent in guiding users from the start of play to the end of tasks, only 8.2% of apps coupled pre-play modeling with post-response scaffolded feedback.

Reward systems. Very few apps (3.5%) provided rewards that might be considered “educational”, such as leveling up of characters to give them new abilities based on past successes to impact how play progresses (e.g., a jet pack reward allows the character to collect numbers at a faster rate, while the number sequences also grow more complex). About half of the apps (49.7%) used rewards, with 31% of those rewards being a prize (e.g., badges, stickers, etc.) for correct answers. Only 1.8% of apps leveled up characters and an additional 1.8% of apps unlocked new levels based on successful progress. Reward systems of unlocking new levels and/or leveling up did not statistically differ between math and literacy apps [$\chi^2 (2, N=171) = 3.37, p=0.19$], nor between Apple and Android platforms [$\chi^2 (1, N=171) = 0.08, p=0.78$]. However, premium apps used prizes as rewards more often (42%) than the free (21%) apps [$\chi^2 (1, N=171) = 9.22, p<0.01$].

Structure of Challenge

Leveling. Subtle leveling was used by some apps (15.2%), while obvious leveling was used by fewer apps (9.9%). Some (11.7%) allowed users to select their preferred level of difficulty.

Scaffolded challenge. Mobile device apps have the ability to adapt the level of challenge based on user input. As such, the occurrence of this was coded by playing each game well and poorly. This was done within puzzles and across puzzles in each app. Overall findings showed that scaffolded challenge was rarely used. Within puzzles, 68.4% of apps did not scaffold challenge based on user input and none of the apps decreased in challenge based on user failure. However, 30% of apps increased in challenge within puzzles. When looking across puzzles, 80.7% of apps did not modify the level of challenge based on performance (i.e., scaffolding), but 17% of apps increased in challenge if a level was successfully completed.

There was a difference in the use of scaffolded challenge by content area. Almost half of the math apps (41.9%) increased in challenge while only 20% of literacy apps increased in challenge, and 26.5% of apps with both math and literacy content used scaffolded challenge [$\chi^2 (2, N=168) = 8.07, p<0.05$]. While only 20.2% of free apps increased/decreased in challenge, 40.8% of premium apps did so, with the difference being statistically significant [$\chi^2 (1, N=168) = 9.06, p<0.01$]. Lastly, Apple apps used such design significantly more often (37.9%) than Android apps (19.7%) [$\chi^2 (2, N=168) = 7.03, p<0.01$].

Additional Mobile App-based Interactions

Touchscreen interactions used. Complex touchscreen interactions were rarely used across the apps coded. Most apps (65%) used dragging, including 74.7% of the Apple apps and 52.6% of the Android apps [$\chi^2 (1, N=171) = 9.06, p<0.01$].

Some apps (27%) used tracing along the screen to complete tasks, but these did not differ by platform. A few Apple apps (2.3%) incorporated multi-touch (more than one finger taps or drags), with a statistically significant difference between Apple (4 apps utilized multi-touch) and Android (no apps utilized multi-touch) platforms [$\chi^2 (1, N=171) = 3.28, p=0.07$]. None of the apps used the pinching motion to zoom in and out of images. There were no differences between math and literacy apps or free and paid apps.

Mobile-device capabilities. Tablets have built-in functions, such as tilting (accelerometer) and camera-use, that could provide unique learning opportunities for learners, but these functions were rarely observed in the apps reviewed. Only 5 apps (3%) utilized the accelerometer feature, and of those, 4 were Apple and paid. Only 2 apps (1.2%) used the device's camera, both of which were math apps. Lastly, only 4 apps (2.3%) utilized a voice recorder, all of which were paid Apple apps.

In total, across all touchscreen and mobile-device capabilities, Apple apps used significantly more features compared to Android apps [$\chi^2 (1, N=171) = 12.30, p<0.001$].

Discussion

By examining how popular educational preschool apps are designed to teach, we found that these preschool apps utilized many different teaching tactics, but many of the tactics were not optimal in light of what research has found helps preschool-aged children learn. For instance, the use of feedback that could explain failures and how to succeed in play was rarely used. This is unfortunate, as apps could be designed to deconstruct math and/or literacy material, helping young children to deeply understand the content, rather than just guess in hopes of finding the correct answers (Gavriushenko, Karilainen, & Kankaanranta, 2015). Interestingly, many of the apps examined solely utilized feedback to praise users and/or prompt them to “try again”. While

this may be a commonly used design feature to increase morale of users as they progress through game play, researchers suggest that this may in fact undermine users' intrinsic motivation (Deci, Koestner, & Ryan, 2001; Zichermann & Cunningham, 2011). This form of feedback may allow users to learn through trial and error (Blair 2013), but research finds that it is not a useful teaching tactic for allowing learners to deconstruct and reflect on the educational content (Wooters & Oostendrop, 2013).

Moreover, many of the apps examined did not utilize adaptive challenge that increased throughout successful game play and reduced during failures. There is a degree of learning that a person can achieve when working independently but providing adaptive challenge can help learners to transcend beyond their independent level of learning to a more complex and sophisticated level (Vygotsky, 1978). From this review, it appears that guided challenge could be better applied to the design of educational apps for preschool learners (Masterman & Rogers, 2002; Sharma & Hannafin, 2007) with the intent to foster deeper understanding of material (Graesser, McNamara, & VanLehn, 2005).

Furthermore, while some apps used descriptive prompts that could help preschoolers understand the goals of each task, few apps utilized modeling to demonstrate how users could get to correct answers and/or avoid incorrect answers. Young children can learn from well-designed instruction and modeling (Kloos & Orden, 2005), but without such guidance, children are left to learn the game's goals solely through trial-and-error. With too many distractors, young children are often unable to sustain attention (Kannass & Colombo, 2007) and could lose focus of the goal(s) of the task (Marcovitch et al, 2007). Modeling could be used to clearly show a child the steps needed to correctly complete tasks (Kloos & Orden, 2005).

With that said, there were some features of the examined educational apps that did align well with preschooler development capabilities. For example, most apps provided prompts verbally, with a minimal use of textual prompts and feedback across apps, which can support an audience of emerging readers (Lonigan, Burgess, & Anthony, 2000). Furthermore, touchscreen interactions commonly utilized by apps were rarely complex, requiring basic hand motions that were appropriate for preschoolers' developing fine motor skills (Vatavu, Cramariuc, & Schipor, 2015). While these features appear in support of preschool learning, many more features could be improved to truly align with evidence-based teaching practices for preschoolers.

Interestingly, there were few differences between apps sold on Apple and Android markets, whether they were free or premium, or targeted math or literacy. However, premium Apple apps tended to show slightly more of these evidence-based teaching practices by scaffolding challenge (but not feedback). Math apps were also more likely to scaffold challenge than literacy apps. Premium Apple apps were also more likely to use unique tablet features, though these features were relatively uncommon across all apps.

Limitations & Future Directions

It is important to note the limitations to this study. First, we did not observe children playing these games and their subsequent learning. This is a content analysis based exclusively on the presence or absence of research-supported teaching practices for these ages. Second, it is possible that our framework missed other important design features. Third, although games were played repetitively, in many different ways, across several coders, it is possible we missed some features in specific apps. Fourth, only the “top” apps advertised on Apple and Android markets were selected. There are many more math and literacy apps on the market, and some of them may have educational features that account well for many developmental capabilities of

preschool-aged learners. However, by targeting the top apps, we reviewed the games to which most children have access. Future research should examine the design of more apps, perhaps randomly selected, to deconstruct how they are designed and experimentally test how certain design features may be more effective for teaching preschool-aged learners than others. Differences between app markets, mobile device platforms, and app costs could also be further compared and even tested to see how they impact user interest, interactions, and learning.

Conclusion

A large percentage of mobile devices are accessible to families (Plowman, McPake, & Stephen, 2010; Pew Research Center, 2017) and educational apps are often advertised for preschool-aged learners (Shuler, 2012). This study details how these widely accessible apps that are marketed as educational are designed in light of extant knowledge of how young children learn best. Understanding how popular preschool apps are designed creates opportunities for understanding how apps may or may not be connected to existing literature on young children's learning, which could inform the design of future apps. App developers should consider how certain designs for prompting, feedback, leveling, and rewards may best support young preschool learners. Moreover, this review could help parents, caregivers, and educators, who want to use educational apps for teaching young children, identify important features when selecting apps. Ultimately, these findings help us understand design features commonly and rarely used in the current educational preschool app market, which can assist app market consumers, as well as inform designers of math and literacy apps for preschool-aged children

CHAPTER 3: STUDY 2

Which App Feedback and Leveling Designs Best Scaffold Preschool Learning?

For years, educators and researchers have strived to identify which teaching practices are most effective for preschool-age learners (e.g., Early et al, 2007; Pianta et al, 2009). With preschool years being foundational for later development and academic success (Campbell et al, 2002), there is little reason to wonder why so much time and energy has been put toward bettering early childhood education (Egert, Fukkink, & Eckhardt, 2018). Moreover, today, as society becomes more reliant on and drawn to digital technologies (Molnar & Cavanagh, 2013), new digital educational tools are being used and more continue to be developed, such as mobile applications (apps) (Shuler, 2012; Vaala et al, 2015). This means that there is now a need for ensuring educational apps are designed to effectively teach young users. Fortunately, this digital transition has positive implications. For example, since most households have mobile devices and apps are free or inexpensive, educational apps are now easily accessible to a wide range of families from diverse socio-economic backgrounds (Pew Research Center, 2015). However, to truly reach apps' highest potential for educating, it is important to determine which educational app designs are most effective for teaching preschool-age learners.

By building on past research examining preschool learning in face-to-face and digital contexts, educational apps can incorporate traditional evidence-based tactics into game design to support preschool learning. One way to do this is by scaffolding; that is, providing preschool learners with hints and support to guide them toward successfully understanding and completing educational tasks (Vygotsky, 1978). Studies show that scaffolding is a successful teaching tactic for helping young children learn and deeply understand new concepts (Hsin & Wu, 2011). Apps can be programmed to scaffold learning in many ways, including well-designed feedback and

leveling. Feedback is an app's response to user-input, informing the user of whether they are progressing through tasks well or if they have made errors. Past research shows that feedback in traditional classroom settings is particularly useful if it explains why certain outcomes have occurred, helping learners to draw connections between their actions and results (Wooters & Oostendrop, 2013). An app's feedback can similarly use this tactic to teach children why certain user-responses are correct and why others are incorrect. Additionally, apps' scaffolding can be further strengthened with strategic leveling that gradually grows in complexity. Building from easy to more challenging content can allow a learner to build basic skill sets in tasks that are foundational for completing later, more difficult tasks (Zichermann & Cunningham, 2011). While, in theory, scaffolding through app feedback and leveling seems logical, it is possible that young children learn differently when being taught via digital devices, compared to in-person contexts.

Therefore, this current study aims to experimentally test whether scaffolding types of educational app designs promote preschool learning. In better understanding how young children effectively learn from digital devices, we can ensure that the educational apps they play today and in the future are actually designed to teach them. The first step toward achieving this goal is understanding whether children learn from digital apps similarly to in-person settings. Using what previous developmental and learning sciences research has identified as important for early childhood learning (e.g., Hirsh-Pasek et al, 2015), this study specifically tests whether feedback and leveling features that are designed with the intent of scaffolding learning are most effective in teaching young children. The following research questions are examined:

1. Which type(s) of app feedback best supports preschool learning?

2. Does an educational app with scaffolded leveling support preschool learning better than the same app with random leveling?

Literature Review

Traditional In-Person Scaffolding

The core purpose of scaffolding learning is to help novices develop skills beyond what they would achieve without assistance, gradually removing guidance, and increasing challenge as learners become more competent (Hsin & Wu, 2011; Obikwelu, Read, & Sin, 2012; Vygotsky, 1978). In face-to-face learning, scaffolding typically begins with an expert guiding a novice learner and having the amount of assistance gradually taper off as the novice learner becomes more competent. Young children are able to learn better as they are challenged through scaffolded teaching (Hsin & Wu, 2011) and such scaffolding can be provided in multiple ways.

The verbal guidance young children receive during educational tasks can scaffold their learning experience. Hammond and colleagues conducted a 3-year longitudinal study that examined the effects of parents' scaffolding with guiding questions and suggestions on their children working on problem-solving tasks at ages 2, 3, and 4 years of age. They found that parents' scaffolding was predictive of children's executive functioning skills (measured using multiple tasks such as shape stroop, spatial span, and reverse categorization) directly (from age 2 to 3 years) and indirectly (from age 2 to 4 years via verbal ability). Similar studies, such as Hughes and Ensor's (2009) longitudinal study of two- to four-year-olds, further support the importance of maternal scaffolding, linking mothers' verbal interactions (e.g., open-ended questions, praise, and encouragement) to preschoolers' emerging verbal and executive functioning skills.

Increased challenge is another form of scaffolding that can push young novice learners'

boundaries to more advanced levels of learning. For example, Hsin and Wu (2011) conducted an experiment with preschool-aged children (ages four and five) to determine how scaffolding from teachers and hands-on materials could assist children in learning about buoyancy. They found that with scaffolded guidance from teachers, children could receive the feedback and challenge they needed to learn more than the non-scaffolded control group (Hsin & Wu, 2011). Studies like these suggest that purposeful, guided learning can support early childhood learning. When mediating learning through the use of technology, scaffolding assistance in educational apps can be structured in a similar manner (Masterman & Rogers, 2002).

Scaffolding with Digital Devices

Feedback. Feedback is an educational app's way of indicating whether a user has completed a task correctly or incorrectly and it can be provided in many different formats – visually, textually, and verbally. For example, visuals can be utilized to demonstrate scientific concepts (e.g., principles of physics), while also providing visual feedback of correct and incorrect actions. In the popularly played Angry Birds app, the game visually demonstrates how different positions of a slingshot can impact the trajectories of flying objects. Herodotou (2016) found that Angry Birds was particularly useful for increasing 5-year olds' understanding of projectiles. The visual representations of motion and trajectory were found to guide preschoolers in understanding physical concepts, providing the proper guidance to see why past attempts were unsuccessful and why successful attempts were made possible.

Though verbal app feedback is popular, studies examining the design of educational preschool games online and on mobile devices show that scaffolded verbal feedback is one of the rarest feedback tactics used (Callaghan & Reich, under review; Nikoyalev et al, under review). This is similar to what researchers have found on the mobile educational preschool app market

(Highfield & Goodwin, 2013; Vaala et al, 2015). Educational children's apps often use drill and practice tasks, with little or no feedback to inform children whether they did something correctly or not (Blair, 2013; Callaghan & Reich, under review; Highfield & Goodwin, 2013). By describing cause and effect, a child is then able to draw connections between what they just did, whether it is right or wrong, and why (Wooters & Oostendrop, 2013; Hsin & Wu, 2011). Yet, limited research has explored how this type of scaffolded feedback in digital play may support preschool-age learning. In fact, to my knowledge, there are no experimental assessments of preschool learning when playing apps with manipulated verbal feedback. Another educational app feature that can scaffold children's learning but has been rarely researched is leveling.

Leveling. Researchers suggest that it is important to strategically design educational tasks to really guide learners to understand the educational content little by little (Obikwelu, Read, & Sim, 2012). One way to do this is by gradually increasing challenge, allowing a user to first build the skill sets necessary to complete easier, more basic tasks, and then gradually building up to later, more complex tasks (Masterman & Rogers, 2002).

Though leveling research with preschool-age learners is limited, studies with older children (elementary through high school) of virtual tutors (artificial intelligence systems that assist with schoolwork) have found scaffolding to be an important and effective component of children's learning (e.g., Ward, Cole, Bolaños, Buchenroth-Martin, Svirsky, & Weston, 2013; Park & Kim, 2015). Mirroring what has been found to be effective in virtual tutoring, leveling up in educational apps could be useful in scaffolding preschoolers through less complex levels to more challenging levels. Doing so may enable games to support the growing skills of preschoolers, just as tapered scaffolding from educators and caregivers (Hsin & Wu, 2011; Henderson, Many, Wellborn, & Ward, 2002) has been effectively used to build upon

preschoolers’ developmental abilities in the past (Hammond, Muller, Carpendale, Bibok, & Liebermann-Finestone, 2012). However, without experimental research to specifically examine how educational app leveling impacts young children’s learning, it is difficult to know whether gradually increasing challenge in apps is effective at teaching preschoolers. This study aims to fill in such gaps of knowledge, to understand how educational preschool apps may best scaffold young learners through strategic feedback and leveling designs.

Method

Participants

There were 240 four- and five-year-old participants recruited from southern California preschools or community events for this study. Half of the participants were female (n=120), with an even gender split between the four-year-olds (male n=81, female n=81) and five-year-olds (male n=39, female n=39). The majority of participants were Hispanic/Latino (61%) and/or from low-income backgrounds (66%). Many participants came from low-income households that did not speak English (33%) or were bilingual (30%). Participants’ parents varied in their highest level of education, ranging from never attending high school (15%), completing high school (19%), taking some college courses (14%), completing a two- or four-year college degree (20%), or completing a graduate/professional degree (10%). Most participants had previous experience using mobile devices (92%) and previously played educational apps (83%). See Table 3.1 for more details about participants’ demographics.

Table 3.1. Participant Demographics

Characteristic	N	%
Gender		
Female	120	50%
Male	120	50%

Race		
American Indian/Alaskan Native	4	1.7%
Asian/Asian American	13	5.4%
Black/African American	4	1.7%
Hispanic/Latino	146	60.8%
Native Hawaiian/Pacific Island	1	0.4%
White/Caucasian	32	13.3%
Multi-Racial	30	12.5%
No Response	10	4.2%
Main Language		
English	73	30.4%
Bilingual	87	36.3%
Non-English	80	33.3%
Parent Highest Level of Education		
No High School	36	15.0%
Some High School	26	10.8%
High School Graduate	45	18.8%
Some College	33	13.8%
Two-year College Degree	17	7.1%
Four-year College Degree	31	12.9%
Some Graduate/Professional School	9	3.8%
Graduate/Professional Degree	25	10.4%
No Response	18	7.5%
Yearly Household Income		
\$0 to \$24,999	105	43.8%
\$25,000 to \$49,999	52	21.7%
\$50,000 to \$74,999	15	6.3%
\$75,000 to \$124,999	10	4.2%
\$125,000 to \$199,999	18	7.5%
\$200,000 +	12	5.0%
No Response	28	11.7%

Procedure

To recruit participants for the study, members of the research team spoke directly with parents/caregivers and preschool directors, describing the purpose of the study and the 10-minute tablet game their child(ren) would play. Parents/caregivers then completed a consent form for their child's participation, along with a questionnaire about their child's background and media use (see Appendix C). Once parents provided written consent, the research team then asked each

child if they were interested in playing the game and brought out the tablet to play the game (e.g., during class-time, at the community event). Only children who provided verbal consent played the game. To avoid distractions from surrounding children and to ensure that children did not overhear the game before it was their turn to play, each child wore headphones during play. All materials and procedures were approved by this University's Institutional Review Board.

There were two parts of the game, creating a 3x2 between-subject design. One part manipulated three types of audible feedback heard, while the other part manipulated two types of leveling progression played. Each participant was randomly assigned to one of the three feedback conditions, creating three groups of 80 children. Each participant was also randomly assigned to one of the two leveling conditions, creating two groups (with the same participants) of 120 children. To reduce the risk of one part of the game influencing the way children played the second part of the game, these conditions were counterbalanced, where half of all participants first played the feedback tasks before the leveling tasks, while the other half of participants first played the leveling tasks before the feedback tasks (displayed in Table 3.2).

Table 3.2. Group Counterbalancing

Group Number	First Tasks Played	Second Tasks Played	Participants per group
1	Feedback Condition 1 (Non-verbal)	Leveling Condition 1 (Scaffolding)	20
2	Feedback Condition 2 (Basic-verbal)	Leveling Condition 1 (Scaffolding)	20
3	Feedback Condition 3 (Scaffolding)	Leveling Condition 1 (Scaffolding)	20
4	Feedback Condition 1 (Non-verbal)	Leveling Condition 2 (Random Order)	20
5	Feedback Condition 2 (Basic-verbal)	Leveling Condition 2 (Random Order)	20
6	Feedback Condition 3 (Scaffolding)	Leveling Condition 2 (Random Order)	20
7	Leveling Condition 1 (Scaffolding)	Feedback Condition 1 (Non-verbal)	20

8	Leveling Condition 1 (Scaffolding)	Feedback Condition 2 (Basic-verbal)	20
9	Leveling Condition 1 (Scaffolding)	Feedback Condition 3 (Scaffolding)	20
10	Leveling Condition 2 (Random Order)	Feedback Condition 1 (Non-verbal)	20
11	Leveling Condition 2 (Random Order)	Feedback Condition 2 (Basic-verbal)	20
12	Leveling Condition 2 (Random Order)	Feedback Condition 3 (Scaffolding)	20
			N = 240

Tablet Game

All participants played a novel vocabulary and sorting game I designed and programmed specifically for this study. The primary goal of the game was to learn the novel names of differently colored and shaped aliens and sort the aliens onto a spaceship. At the start of the game, children were introduced to aliens, giving them the opportunity to learn aliens' novel names and their associated physical features. For example, a child would see Figure 1, and hear: "This is a Froodle. You can tell because Froodles are orange and shaped like a square. That means they have four corners! Drag the Froodle to the spaceship." Children then completed a training task in which they needed to drag the alien to the spaceship. Once successful, the second training phase started in which the scene transitioned to a new alien with another novel name, such as a Plickatoo. This task not only allowed the research team to see that the child could identify which object was the Froodle, but to also show that they knew where the spaceship was. Such knowledge was key to be able to complete the sorting tasks to follow.

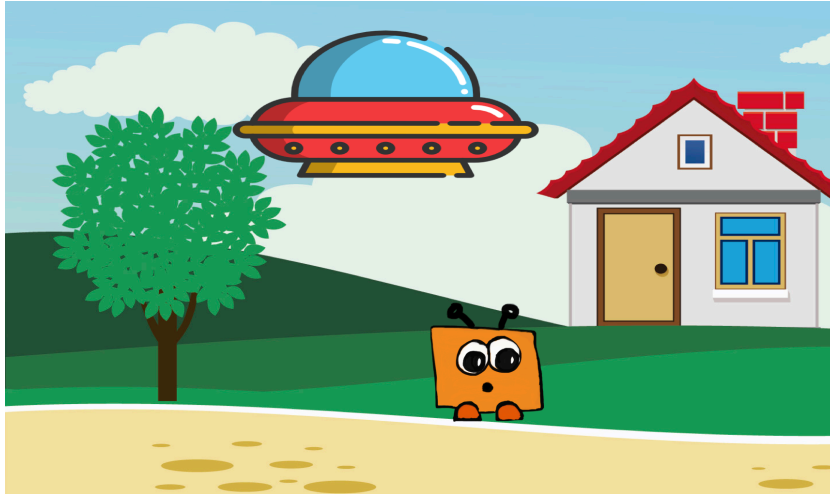


Figure 3.1. Frooodle Introduction Scene

After alien introductions, children were then taken to a screen with multiple types of aliens scattered around, and asked to sort specific aliens, such as the Frooodles, onto the spaceship (see Figure 2).



Figure 3.2. Sorting Task Example

This game was intentionally designed to be played in a single, 10-minute session, so the learning outcome was something that children could quickly grasp. Since young children are often skilled at making associations between new words and their meanings quickly, also referred to as fast mapping (Heibeck & Markman, 1987), novel words were an achievable

learning goal for short-term game play. See Appendix D for a list of all novel alien names, prompts, and responses from the Feedback and Leveling tasks.

Feedback Design. In the feedback portion of the game, children were introduced to four categories of aliens and required, in proceeding scenes, to correctly sort three of them as they sat alongside other distractor aliens. During the sorting tasks, every time a child responded correctly or incorrectly, they heard a different type of audible feedback, depending on which experimental condition they were assigned. This was a between-subject, three-group design because there were three types of audible feedback tested – non-verbal (sound effects), basic verbal (e.g., “great job” or “try again”), and scaffolded verbal (e.g., “That’s a Plickatoo because it’s green!”) (see Table 3.3). All feedback tasks were designed to be equally challenging, to avoid unforeseen confounding effects of leveling.

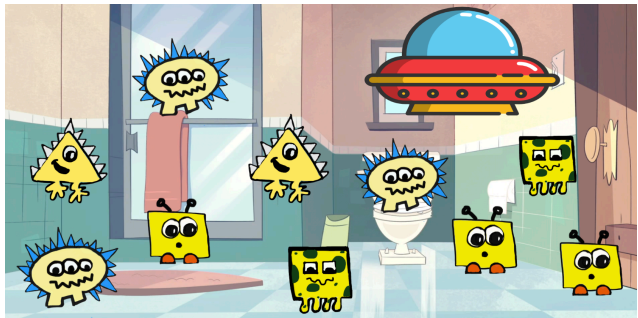
Table 3.3. Examples of Feedback Types

	When correct	When incorrect
Non-verbal feedback (sound effects)	Bell rings	Horn honks
Basic verbal feedback	“Great job!”	“Oops, try again!”
Scaffolded verbal feedback	“That’s a Plickatoo! It’s green.”	“Remember, Plickatoos are green.”

Leveling Design. There were two between-subject conditions for the leveling tasks, with children assigned to random order (a non-strategic sequence of challenge) and scaffolded order (a gradual increase in challenge) leveling. Challenge was manipulated by the types of aliens displayed on the screen during each task. Easier tasks included a group of aliens with all different colors and shapes, while the more challenging tasks had aliens of the same color and similar shapes. An example of a scaffolded condition task is in Table 6 below. The tasks within both conditions were identical in every way, except for their order, therefore, the participants in

the random order condition completed the same exact tasks as displayed in Table 3.4, but in a pre-determined random order. To minimize the chances of audible feedback influencing learning outcomes for the leveling tasks, simple sound effects were used for both leveling conditions, solely to indicate whether responses were correct or incorrect.

Table 3.4. Scaffolded Leveling Design

Scaffolded Challenge to find Yooks 	
Task Prompt	Visuals
“Do you see all of the Yooks in this room? We need your help moving them to the spaceship. Drag the Yooks to this spaceship!”	
“There are still some Yooks in this room! Put all of the Yooks on the spaceship!”	
“Here’s another room with aliens. Please put the Yooks on the spaceship!”	

Measures

Game data: Errors and time. The dependent variables measured in this study were participants' number of incorrect actions (i.e., errors) and their response times across all scenes (3 aliens for feedback and 3 aliens for leveling, 3 scenes per alien = 6 aliens and 18 scenes total). The app was programmed to record number of errors and response times per task during game play. Learning was operationalized by these measures, where fewer errors represented participants' accuracy and lower response times indicated faster completion. Thus, smaller values on both of these outcomes indicate greater success in the game play.

Background data. The demographics questionnaire given to parents was used to gauge participants' experiences using technology, especially tablet technology, to account for any group similarities or differences in technological skills that may have impacted their performance in the experiment. This questionnaire was also used to measure demographic information such as participants' age, ethnic background, family income, language, and parental education.

Analysis

T-tests and analysis of covariance (ANOVA) with Tukey post-hoc tests were used to analyze between-group differences for the three feedback conditions and two leveling conditions. The independent variables were the three feedback conditions and the two leveling conditions. The dependent variables (DV) were the number of incorrect responses (errors) and the amount of time it took participants to complete each task (response times). These DVs were used in separate analyses.

The covariates included within the models were: child's age, previous tablet use, and the order of play (whether they played the feedback tasks or leveling tasks first). Research shows that children improve their touchscreen capabilities with tablet experience and increased age

(Vatavu et al, 2015). Therefore, age and prior tablet use were both controlled for in these analyses. This tablet use was categorized into three levels – no experience, little experience (up to 2 days a week), much experience (up to seven days a week). Since the feedback and leveling tasks were counterbalanced, the order of play was also controlled for; to account for any training effects prior tasks would have on the second half of game play.

Results

Feedback Design

Number of Errors. First, analyses were done to examine how children in the feedback groups differed in their number of errors made. An ANCOVA test revealed that there was a main effect of age [$F(1,240)=6.11, p<0.05$], but the three feedback groups did not statistically differ from one another. However, the group means (displayed in Table 3.5) and low p-value suggest that the data were trending toward statistical significance. To examine this further, a second ANCOVA was run to see if an effect of feedback was present for those with no training on the alien game (i.e., only those with the feedback condition first). This second ANCOVA revealed that the children in the scaffolded feedback group ($M=1.71, SD=1.22$) made significantly fewer errors than the children in the sounds feedback group ($M=1.12, SD=1.12$); [$F(2,120)=6.20, p<0.01$].

Table 3.5. Means and Standard Deviations of Participants' Errors in Feedback Tasks

Age Group	Feedback Group	<i>n</i>	<i>Errors in Feedback Tasks</i>	
			<i>M</i>	<i>SD</i>
Both	Group 1 (Sounds)	80	1.48	1.33
	Group 2 (Encouragement)	80	1.20	0.88
	Group 3 (Scaffolded)	80	1.11	0.92

4	Group 1 (Sounds)	55	1.55	1.29
	Group 2 (Encouragement)	55	1.28	0.98
	Group 3 (Scaffolded)	52	1.30	1.09
5	Group 1 (Sounds)	25	1.28	1.17
	Group 2 (Encouragement)	25	1.01	0.58
	Group 3 (Scaffolded)	28	0.77	0.46

Response Times. When looking at the main effect of feedback conditions on response times, an ANCOVA test showed that there was a statistically significant difference between the feedback groups. On average, the sounds feedback group ($M=11.13$, $SD=4.86$) completed tasks statistically significantly faster than the encouragement feedback group ($M=13.65$, $SD=8.64$); [$F(1, 240)=3.44$, $p<0.05$]. There was also a main effect of age ($p<0.001$; see Table 3.6 below) and order of play ($p<0.001$). However, unlike accuracy, the effect of scaffolded feedback on response times for participants who played the feedback tasks first ($n=120$) did not statistically significantly differ.

Table 3.6. Means and Standard Deviations of Participants' Response Times for Feedback Tasks

Age Group	Feedback Group	<i>n</i>	<i>Response Times for Feedback Tasks</i>	
			<i>M</i>	<i>SD</i>
Both	Group 1 (Sounds)	80	11.13	4.86
	Group 2 (Encouragement)	80	13.65	8.64
	Group 3 (Scaffolded)	80	12.16	5.94
4	Group 1 (Sounds)	55	11.72	5.15
	Group 2 (Encouragement)	55	15.04	9.87
	Group 3 (Scaffolded)	52	13.39	6.79
5	Group 1 (Sounds)	25	9.84	3.97
	Group 2 (Encouragement)	25	10.58	3.47
	Group 3 (Scaffolded)	28	9.88	2.80

Leveling Design

Number of Errors. When comparing scaffolded leveling with random order leveling conditions, a t-test was first used to determine if there was a statistically significant difference between the two conditions. Results showed that this main effect of leveling was statistically significant ($t(238)=-3.17, p<0.001$), even when including covariates in an ANCOVA test; [$F(1, 240)=10.52, p<0.01$]. A post-hoc analysis showed that the scaffolded leveling group ($M=2.54, SD=1.96$) had fewer average errors than the random order leveling group ($M=3.34, SD=1.96$). There was no statistically significant main effect of age or tablet experience between the two groups, but there was for order of play ($p<0.001$). See Table 3.7 below.

Table 3.7. Means and Standard Deviations of Participants' Number of Errors of Leveling Tasks

Age Group	Feedback Group	<i>n</i>	<i>Number of Errors for Leveling Tasks</i>	
			<i>M</i>	<i>SD</i>
Both	Group 1 (Scaffolded)	120	2.54	1.96
	Group 2 (Random Order)	120	3.34	1.96
4	Group 1 (Scaffolded)	83	2.63	2.10
	Group 2 (Random Order)	79	3.48	2.00
5	Group 1 (Scaffolded)	37	2.32	1.59
	Group 2 (Random Order)	41	3.06	1.86

Response Times. A t-test was used to examine differences between the response times of the two leveling conditions and it showed that there was a statistically significant difference between scaffolded and random order leveling; [$t(238)=-3.43, p<0.001$]. The scaffolded leveling group ($M=17.18, SD=8.62$) completed tasks faster than the random order leveling group ($M=21.51, SD=10.88$). After running an ANCOVA test, the leveling groups were still

statistically significantly different; [$F(1, 240)=13.92, p<0.001$]. Additionally, there was a statistically significant main effect of age [$F(1,240)=11.19, p<0.001$] and order of play [$F(1,240)=11.02, p<0.01$]. See Table 3.8 for more group details and Appendix E for ANCOVA tables.

Table 3.8. Means and Standard Deviations of Participants' Response Times for Leveling Tasks

Age Group	Feedback Group	<i>n</i>	<i>Response Times for Leveling Tasks</i>	
			<i>M</i>	<i>SD</i>
Both	Group 1 (Scaffolded)	120	17.17	8.62
	Group 2 (Random Order)	120	21.51	10.88
4	Group 1 (Scaffolded)	83	18.28	9.25
	Group 2 (Random Order)	79	23.43	12.17
5	Group 1 (Scaffolded)	37	14.67	6.41
	Group 2 (Random Order)	41	17.81	6.50

Discussion

This study tested whether scaffolded verbal feedback (explaining right or wrong responses) and scaffolded leveling (gradually guiding learners toward more difficult tasks), could improve children's learning of novel words from an app. Findings suggested that for novice players, participants' accuracy in selecting correct answers could be better with scaffolded verbal feedback that describes right or wrong answers, compared to sound effects feedback. The absence of a significant difference when using the entire sample suggests that the effect, if present, is relatively small – thus warranting a larger sample for sufficient statistical power to detect the effect. However, the presence of an effect with half the sample suggests that when tasks are simple and novel, scaffolding feedback helps. However, after sufficient trial and error play (i.e., leveling game), children no longer benefit from the scaffolded feedback. It is possible

that through play, children identify the important features that scaffolding would have highlighted. Still, additional studies on the design of scaffolded feedback are needed. The finding that scaffolded feedback can improve accuracy aligns with previous studies which show that scaffolding can increase young children's learning (Hammond et al, 2012; Herodotou, 2016; Hsin & Wu, 2011).

Additionally, though research shows how young children benefit from digital verbal praise (Bracken & Lombard, 2004), encouraging feedback actually slowed participants down, compared to sound effects feedback. Yet, other researchers have also argued, that praise and extrinsic rewards can undermine motivation under certain circumstances (Deci, Koestner, & Ryan, 1999; Henderlong & Lepper, 2002), which may have been why participants in this study had longer response times. Also, the excitement associated with being praised may have led participants to briefly celebrate the success, before attempting to complete the next tasks. Perhaps a solution to maintaining intrinsic motivation and faster responses to tasks is coupling feedback which praises and scaffolds children (e.g., "great job!" and "oops, try again!"), thus providing encouragement and instructions.

Another way of increasing accuracy and speed is through scaffolded leveling designs. Findings showed that gradually increasing the challenge helped participants get fewer incorrect answers, compared to the participants who played the app with a random order of leveling. Guiding children to first understand easy, foundational concepts, can better prepare them to be more accurate during their later, more challenging tasks (Obikwelu, Read, & Sim, 2012). These finding show how apps can scaffold children's learning from basic to intermediate content. Further testing showed that scaffolded leveling was also associated with lower response times. Future research can explore whether adaptive leveling (decreasing challenge after incorrect

answers and increasing challenge after correct answers) can have an even stronger positive effect on children's learning. Meeting children at their personalized level of development and learning potential (Vygotsky, 1978) may be what educational apps need to best guide children. By utilizing such scaffolded feedback and leveling, educational apps can become the effective, easily accessible tools they are intended to be.

Limitations

There are a few limitations of this study that should be considered. First, about 68% of the analysis sample was four-year-old participants. Because there were some differences in the effects of game designs between four-and five-year-old participants, collecting data from more five-year-old participants may provide more statistical power for comparing and detecting smaller effects. However, research shows that age can impact children's performance on cognitive tasks (Mahy, Moses, Kliegel, 2014), so age differences identified are consistent with research showing that children process information differently within the short span of preschool-years.

Additionally, this study solely looked at the effects of feedback and leveling in isolation from one another. Typically, educational preschool apps incorporate both feedback and leveling features simultaneously. Many face-to-face studies have shown that the combination of guiding hints and reduced assistance to increase challenge over time, is effective for teaching preschool-age learners (Hsin & Wu, 2011). Therefore, future studies can examine whether the combination of scaffolded feedback and scaffolded leveling similarly guides young children to learn. Such studies that integrate feedback and leveling would not need to counterbalance, thus enabling researchers to see the direct effects of the manipulated app designs, without child learning outcomes being influenced by potential training effects.

Lastly, this study examined an app that was solely used once by participants, for a brief amount of time, and measures did not include long-term or transfer effects. Future research can examine whether scaffolded designs have a lasting effect on preschool learning and whether it is more beneficial for transferring learned knowledge/skills to out-of-game contexts. In doing so, educational preschool apps can be designed to educate young children for long-term learning that could eventually transfer beyond the tablet app interaction.

Conclusion

Educational apps are popular platforms for early childhood learning in this digital age (Highfield & Goodwin, 2013) and the intent of this study is to assist game designers in strategically crafting effective educational games to fit the abilities and desires of preschool-aged children. Scaffolding has been an effective teaching tactic used for years in face-to-face contexts (e.g., Hsin & Wu, 2011), but now there is a need for discovering how digital tools can be designed to support young learners in the same way. Effective educational apps can impact preschool learning within games and in out-of-game contexts (Huber et al, 2016; Herodotou, 2016), but more research, like this study, is needed to identify which features are truly effective in scaffolding young learners. Doing so can provide a framework of overarching, effective design features that can be utilized across the plethora of educational preschool apps on the market today and in the future.

CHAPTER 4: CONCLUSION

How Can Digital Apps Support Early Childhood Learning?

Important Contributions

Educational preschool apps have the potential to provide engaging learning experiences to young children from a broad range of backgrounds, locations, and socioeconomic situations, yet researchers have only just begun to examine how those apps are designed to teach and which designs are most effective for early childhood learning. This dissertation provided several important contributions to current empirical literature, demonstrating how developmental and educational research can be used as a foundation for designing digital learning tools for young children.

Study 1 extended prior research that looked at the overarching designs of educational preschool apps (Highfield & Goodwin, 2013) by examining many, if not all, of the features children come across during game play – including styles of prompting, character modeling, formats of feedback, leveling designs, and reward systems. Through this extensive examination of the design features utilized by educational preschool apps, using a framework specifically created to address developmental evidence-based teaching practices, Study 1 provided a wholistic, developmentally-appropriate measure of educational preschool app designs. This was crucial to conduct because early childhood is a significant time in development (Campbell et al, 2002), making it very important to ensure the educational tools children use are actually designed to capitalize on their learning capabilities (Hirsh-Pasek et al, 2015).

While empirical findings from face-to-face studies in developmental and educational research are useful in understanding how young children may learn from educational apps, experimental testing is ultimately required to determine if in-person tactics are just as effective

for teaching preschoolers via digital devices. Study 2 contributed to existing literature by experimentally testing the direct effects of certain educational app designs to deduce how these apps can effectively scaffold preschool-age learners. Prior research has examined some designs of feedback (Blair, 2013), but Study 2 is the first to isolate and manipulate the effects of audible feedback designs. With audible feedback being one of the most commonly used forms of feedback (Callaghan & Reich, under review), it is essential to understand which types of audible feedback are best for scaffolding young learners. Moreover, with so few preschool apps designed to utilize challenge in a way that gradually guides children toward more complex content (Callaghan & Reich, under review), Study 2 findings can be utilized to show app developers and consumers just how important scaffolding challenge is for preschool-age learning.

Overall, this dissertation aimed to improve the design of educational preschool apps to be sold on our increasingly popular app markets. This dissertation's content analysis shows that educational preschool apps have much room for improvement, to align with evidence-based teaching practices, and the experimental findings shows that digital early childhood learning can be scaffolded with guiding designs. The following subsections will show what we have to learn from these two dissertation studies, showing the importance in understanding these digital tools that are readily available in most children's households (Pew Research Center, 2017). The first step toward understanding these apps is by examining the designs of educational preschool apps on the app markets today.

Chapter 1: Main Points

To identify whether today's popular educational preschool apps were designed with the developmental abilities of young learners in mind, Study 1 provided a detailed examination of the top educational math and literacy preschool apps (N=171) on today's Apple and Android

markets. Specifically, Study 1 aimed to answer: How are popular preschool math and literacy apps designed? And, do their features connect with evidence-based teaching practices? There were three main findings from this first study: 1) Apps often give clear goals and are “positively” rewarding, 2) there are several features that can be improved to align with empirical-based teaching practices, and 3) opportunities for harnessing the unique capabilities of tablet technology are frequently missed.

Clarity and Positivity. Today’s popular educational preschool apps are designed with young children’s distractibility in mind by creating clear goals. Further, they aim to provide users with a positive “feel good” experience while they play. However, these designs alone may not be sufficient enough for preschool learning to truly occur. Almost 80% of the apps examined provided clear goals for the learner to understand what is needed to complete the tasks, but after the initial goals were given, little was done to ensure that young learners remain informed and on-task. Of the few studies that have tested learning from apps, modeling appears highly important for prompting preschool-age learners (Hiniker et al, 2015), yet, preschool apps rarely (4.7%) model how to complete tasks correctly. This would require children to solely base their efforts on the goals stated at the start of play. Furthermore, preschool-age children are easily distracted (Kannass & Colombo, 2007), so ensuring that they remain on task is highly important during app play. However, few apps provide in-play guidance of repeating (15.2%) or rephrasing (6.4%) instructions for children who may not have heard or understood the original directions. So, while clarity is made a priority at the very start of tasks, preschool apps seem to lack the guidance needed to truly guide learners to understand the tasks and how to successfully complete them.

Furthermore, preschool apps were very good at providing positive rewards and responses to user-input. While giving a child prizes, like stickers and badges, may make the game fun and rewarding, helping a child feel good during play, researchers suggest that this tactic also undermines learners' motivation for completing educational tasks (Botha, Herselman, & Ford, 2014; Zichermann & Cunningham, 2011). Another "feel good" tactic commonly used by today's preschool apps is positive feedback. At the successful completion of tasks, celebratory visual (68.4%) and verbal (73.7%) feedback was frequently used by preschool apps, and most (46.8%) apps used non-scaffolded verbal feedback after incorrect responses. Researchers have argued that children desire to be praised, but sometimes praise can undermine children's learning and motivation to learn (Henderlong & Lepper, 2002). This dissertation's study 2 experiment found that positive feedback does not increase children's accuracy compared to other feedback types, suggesting that there is little educational benefit when apps solely provide positive verbal feedback. While extrinsic rewards and praise may be useful in building preschoolers' morale, research shows that motivation to learn can be weakened by such rewards (Deci, Koestner, & Ryan, 1999). Preschool apps are very good at giving clear goals and positive responses to users, but they greatly lack in guiding children to deeply learn.

Lack of Scaffolded Designs. There are many ways in which apps can be designed to scaffold preschoolers' learning, but these designs are scarcely used. Research shows that children benefit from being guided toward learning new educational content, through useful tips, guiding questions, and the tapering away of support to gradually create more challenging tasks (Hsin & Wu, 2011). Yet, Study 1's findings showed that these apps are designed as drill and practice, solely indicating whether responses were correct or incorrect, with limited opportunities for deeply understanding the reasons behind certain outcomes; aligning with previous studies

(Highfield & Goodwin, 2013). For example, rather than scaffolding learning by adapting the level of challenge to user-performance or gradually increasing in challenge, most apps (80.7%) do not modify their difficulty. This is unfortunate given that Study 2's experimental finding that scaffolded feedback has the potential to support preschool learning, and gradual increases in challenge are more effective in guiding preschoolers toward accurate and rapid responses. The effective scaffolding tactics for preschool learning are the rarely used tactics in today's educational preschool apps. Much more can be done to create educational preschool apps that utilize a range of features to facilitate deeper learning experiences.

Missed Tablet Opportunities. The unique characteristics and capabilities of tablet devices could also be used to facilitate deeper learning experiences. Tablet devices are innovative tools that can support preschool learning in ways that books and hands-on manipulatives could never do. Yet, it seems that apps are being utilized as though they are solely a flat surface to project educational content on the screen, and instead of using a pen or pencil, learners can interact with their fingertips. However, tablet devices are equipped to do so much more if the features are taken advantage of. For example, when teaching measurement in-person, learners are limited by the resources they have available to test the pouring of different substances into measuring cups and limited to a location where it is acceptable to make a mess with spilt substances. Now, tablet devices could enable children to learn the very same concepts, with the tipping and shaking of the device, with an unlimited amount of measuring substances and the freedom to learn measurements from any location. It may seem that digital devices would take away from the hands-on experiences that children need to truly learn, but it is feasible that, when designed and used creatively, tablet devices could be truly valuable and effective educational tools (Herodotou, 2016; Klahr, Triona, & Williams, 2007). Researchers

have begun to examine such possibilities (Yannier, Koedinger, & Hudson, 2015), and future studies could extend this work by examining the variety of content areas that may be taught with digital devices and how combinations of tablets' unique capabilities may be capitalized on. In doing so, we can better understand whether/how tablet capabilities can best facilitate learning, especially since current popular preschool apps rarely use these features.

Concluding Study 1. Ultimately, there is much room for improving the educational design features of educational preschool apps. Currently, apps are great at praising and encouraging young children as they play, but rarely do apps account for young children's distractibility, limited attention spans, or developing executive functioning skills. In efforts to better understand how young children respond and learn from digital apps, the second study of this dissertation aimed to identify direct effects of certain app designs.

Chapter 2: Main Points

Research shows that preschool-age learners benefit from helpful tips and guiding challenge in-person (Hsin & Wu, 2011), also referred to as scaffolding, but there is little understanding of how digital educational apps may be designed to scaffold preschool learning. Two main design features of apps that can be utilized for scaffolding young children's educational experiences is feedback and leveling. Feedback can be used to help children learn why certain answers are correct and while other answers are incorrect, while leveling can be utilized to guide young learners from easy to more challenging tasks. The second study of this dissertation sought to determine if scaffolded feedback and leveling are most effective for teaching preschool-age learners novel names in a sorting task. Specifically, the following research questions were explored:

- 1) Which type(s) of app feedback best support preschool learning?

2) Does an educational app with scaffolded leveling support preschool learning better than the same app with random leveling?

There were two main findings to this study: 1) Scaffolded feedback may be useful for accuracy, while encouraging feedback leads preschoolers to take longer in completing tasks, and 2) Scaffolded leveling can lead to more accurate and faster task completion.

Feedback for Support. Feedback is an app's way of indicating whether a user completes certain tasks correctly or incorrectly, but feedback can be utilized for teaching the content as well. Study 2 found that when young children are given verbal descriptions for why certain answers were correct and why other answers are incorrect, they can be more accurate in their responses compared to children who are solely given sound effect notification of whether their response was right or wrong. Since this was only the case for children who played the app feedback tasks first, it suggests that the benefit of scaffolded feedback is stronger if that guidance is provided from the very start of play. If other forms of feedback are used first, children may become trained to not utilize feedback to its fullest potential or simply learn from trial and error, making later scaffolding feedback less beneficial. Nevertheless, the positive effect scaffolded feedback had on learners' accuracy aligns well with past research that shows the educational benefit of scaffolding young learners in face-to-face settings (Hsin & Wu, 2011). However, scaffolded feedback was rarely used by popular educational preschool apps, suggesting that the feedback in today's preschool apps are not the most supportive for teaching young learners. More research is needed to understand how to strengthen the effect of scaffolded verbal feedback and how other forms of feedback, such as scaffolded visual feedback, may be utilized to guide children toward grasping the reasons behind certain outcomes and ultimately learn how to avoid making mistakes.

Though scaffolded feedback may be useful for accuracy, Study 2 also found that participants take longer to complete tasks when they have encouraging feedback, compared to children who solely hear sound effects. There are multiple potential explanations for this outcome. One is that children may take more purposeful time to strategize in their selections, in the hopes of receiving more positive feedback (Tarullo, Nayak, St. John, & Doan, 2018). Based on anecdotal observations of the participants playing the app, there were many instances of children celebrating (e.g., shouting “Yes, I did it!”, while throwing hands up with a smile), suggesting a second possible explanation, that children take more time to celebrate their victories after receiving encouragement feedback. Or, perhaps the encouragement feedback leads to a decrease in children’s motivation to persist (Henderlong & Lepper, 2002). Ultimately, the encouragement feedback was not useful in helping children make more accurate selections, compared to their two counterpart groups. Future research can further explore whether the extended response times from encouragement feedback could benefit or hinder children when learning from educational apps.

Leveling for Guidance. Scaffolded leveling is a highly supportive feature for preschoolers to make more accurate and rapid decisions. During the early childhood years of ages three to five, children often vary in their knowledge and skill sets. This makes it challenging to know just what level of sophistication a child is at when they start playing an educational app. By beginning an app with easier tasks and gradually building toward more challenging content, young learners of all skills and learning levels may be able to benefit from the tasks. This is because the new learners are able to establish the basic skill sets needed for later, challenging tasks, while the advanced learners receive extra practice in the basic tasks before moving on to the tasks purposefully designed to challenge them. More research can be used to examine the effects of

scaffolded leveling on younger ages, such as 3-year-olds, to see if that gradual support remains beneficial. Additionally, future studies can also test whether adaptive leveling provides guidance that is just as effective, or perhaps more effective, at meeting young learners at their personal levels of learning. Studies have found that adaptive computer programs have been useful in enhancing children's working memory skills (Holmes, Gathercole, & Dunning, 2009), so perhaps this is possible for learning new content as well.

Recommendations

Findings from this dissertation can be utilized by both educational preschool app developers and consumers. App developers can better consider how young children best learn and how educational apps can be designed to effectively support early childhood learning. This can be done by referring to previous research examining effective face-to-face teaching tactics for young children, further exploring whether certain combinations of feedback designs are more effective and incorporating scaffolded leveling techniques. While it is important for studies like these to help identify overarching design features that are most effective for preschool-age learning, developers should strive to critically test the effectiveness of their individual products as well. Such testing can include factors like children's enjoyment during game play, especially understanding how to disentangle purposeful and playful errors made by children from lack of understanding and/or ways to celebrate successes may be designed as an additional opportunity for learning (Fróes & Tosca, 2018). Ultimately, the educational apps released on the market need to be tested for effectiveness, to ensure they can truly produce higher learning outcomes. Doing so can help ensure that the educational preschool apps on today's app market are genuinely educational for preschool-age learners.

Moreover, educators, parents, and caregivers, who are consumers of these digital tools for their children can utilize this research to be better aware of the types of games they should invest in. By making informed decisions, young children can be handed educational apps that truly support their learning in fun, innovative ways. Our digital society has grown to be rapid and frequent users of apps for mobile devices (Pew Research Center, 2015), and consumers must be cautious about judging an app solely by its description (Shuler, 2012). Simply because apps claim to be educational, does not mean that companies thoroughly tested the effectiveness of their products. Studies like these, help enable caregivers and teachers of preschoolers to be more critical consumers of apps and provide young children with beneficial digital learning tools.

Future Research

Though this dissertation makes some strong contributions to existing literature, there are always future directions for studies to come. Educational preschool apps on the market today have much room for improvement and more experimental research can help create a set of guidelines for designing early childhood apps. More research that isolates various features can help identify their individual effects on learning and this should ultimately lead to experiments that examine the effects of combining app features together. Moreover, these studies should not only examine in-game learning, but incorporate transfer tests as well, to understand how in-game learning can also be useful for out-of-game contexts. Ultimately, since educational preschool apps integrate multiple features (e.g., prompts, feedback, leveling, and rewards) into a single stream of play, identifying which combinations best support preschool age learners can create cheap, effective educational tools that are readily available from the many mobile devices in children's homes (Plowman et al, 2012).

REFERENCES

- Blair, K. (2013). Learning in critter corral: Evaluating three kinds of feedback in a preschool math app. In Proceedings of the 12th International Conference on Interaction Design and Children (pp. 372-375). ACM.
- Botha, A., Herselman, M., & Ford, M. (2014). Gamification beyond badges. In IST-Africa Conference Proceedings, (pp. 1-10). IEEE. Retrieved from:
<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6880651>
- Bracken, C. C. & Lombard, M. (2004). Social presence and children: Praise, intrinsic motivation, and learning with computers. *Journal of Communication*, 54(1), 22-37.
DOI: <https://doi.org/10.1111/j.1460-2466.2004.tb02611.x>
- Callaghan, M. N. & Reich, S. M. (under review). Are Educational Preschool Apps Designed to Teach? An Analysis of the App Market.
- Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project, *Applied Developmental Science*, 6:1, 42-57. DOI: 10.1207/S1532480XADS0601_05
- Carlson, S. M. & Wang, T. (2007). Inhibitory control and emotional regulation in preschool children. *Cognitive Development*, 22(4), 489–510. DOI:
<https://doi.org/10.1016/j.cogdev.2007.08.002>
- Chiasson, S. & Gutwin, C. (2005). Design principles for children's technology. *Interfaces*, 7, 28-37. Retrieved from
http://www.hci.usask.ca/publications/2005/HCI_TR_2005_02_Design.pdf
- Cooper, L. (2005). Developmentally appropriate digital environments for young children. *Library Trends: Project MUSE*, 54(2), 286-302. DOI: 10.1353/lib.2006.0014

- Cuban, L. (2001). *Oversold and underused: Computers in classrooms, 1980–2000*. Cambridge: Harvard University Press
- Deci, E. L., Koestner, R., & Ryan, R.M. (1999). A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin*, 125(6), 627-668. DOI: 10.1037/0033-2909.125.6.627
- Deci, E. L., Koestner, R., & Ryan, R.M. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Educational Research*, 71(1), 1-27. Retrieved from: <http://www.jstor.org/stable/3516064>
- Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, ... Zill, N. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. *Child Development*, 78(2), 558-580. DOI: 169.234.225.23
- Egert, F., Fukkink, R. G., & Eckhardt, A. G. (2018). Impact of in-service professional development programs for early childhood teachers on quality ratings and child outcomes: A meta-analysis, *Review of Educational Research*, 88(3), 401-433. DOI: 10.3102/0034654317751918
- Fróes, I. & Tosca, S. (2018). Playful subversion: Young children and tablet use. *European Journal of Cultural Studies*, 21(1), 39-58. DOI: <https://doi.org/10.1177/1367549417705601>
- Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, 134(1), 31-60. DOI: 10.1037/0033-2909.134.1.31

- Gavriushenko, M., Karilainen, L., & Kankaanranta, M. (2015). Adaptive systems as enablers of feedback in English language learning game-based environments. *Frontiers in Education Conference*. DOI: 10.1109/FIE.2015.7344107
- Gmitrova, V., Podhajecká, M., & Gmitrov, J. (2009). Children's play preferences: implications for the preschool education. *Early Child Development and Care*, 179(3), 339-351. DOI: 10.1080/03004430601101883
- Graesser, A.C., McNamara, D.S., & VanLehn, K. (2005). Scaffolding Deep Comprehension Strategies Through Point & Query, AutoTutor, and iSTART. *Educational Psychologist*, 40(4), 225-234. DOI: 10.1207/s15326985ep4004_4
- Hammond, S. I., Müller, U., Carpendale, J. I. M., Bibok, M. B., & Liebermann-Finestone, D. (2012). The effects of parental scaffolding on preschoolers' executive function. *Developmental Psychology*, 48(1), 271-281. DOI: <http://dx.doi.org/10.1037/a0025519>
- Heibeck, T. H. & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child Development*, 58(4), 1021-1034. DOI: <http://www.jstor.org/stable/1130543>
- Henderlong, J. & Lepper, M. R. (2002). The effects of praise on children's intrinsic motivation: A review and synthesis. *Psychological Bulletin*, 128(5), 774-759. DOI: 10.1037//0033-2909.128.5.774
- Henderson, S., Many, J., Wellborn, H., & Ward, J. (2002). How scaffolding nurtures the development of young children's literacy repertoire: Insiders' and outsiders' collaborative understandings. *Reading Research and Instruction*, 41(4), 309-330. DOI: 10.1080/19388070209558374
- Herodotou, C. (2016). Mobile games and science learning: A comparative study of 4 and 5 years

old playing the game Angry Birds. *British Journal of Educational Technology*, 49(1), 6-16.

DOI: 10.1111/bjet.12546

Highfield, K. & Goodwin, K. (2013). Apps for mathematics learning: A review of ‘educational’ apps from the iTunes App Store. In Steinle, V., Ball, L., & Bardini, C. (Eds.), *Mathematics education: Yesterday, today and tomorrow* (Proceedings of the 36th annual conference of the Mathematics Education Research Group of Australasia). Melbourne, VIC: MERGA

Hirsh-Pasek, K., Zosh, J., Golinkoff, R., Gray, J., Robb, M., & Kaufman, J. (2015). Putting education in “educational” apps: Lessons from the science of learning. *Psychological Science in the Public Interest*, 16(1), 3-34. DOI: 10.1177/1529100615569721

Hiniker, A., Sobel, K., Hong, S., Suh, H., Irish, I., Kim, D., & Kientz, J. A. (2015). Touchscreen prompts for preschoolers: Designing developmentally appropriate techniques for teaching young children to perform gestures. *Proceedings of the 14th International Conference on Interaction Design and Children*, Boston, MA, 109-118. DOI: <http://dx.doi.org/10.1145/2771839.2771851>

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, 12(4), 9-15. DOI: 10.1111/j.1467-7687.2009.00848.x.

Hourcade, J. P. (2007). Interaction design and children. *Human-Computer Interaction*, 1(4), 277-392. DOI: 10.1561/11000000006

Hsieh, H. & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Journal of Qualitative Health Research*, 15(9), 1277-1288. DOI: 10.1177/1049732305276687

Hsin, C. & Wu, H. (2011). Using scaffolding strategies to promote young children’s scientific understandings of floating and sinking. *Journal of Scientific Educational Technology*, 20,

656-666. DOI: 10.1007/s10956-011-9310-7

Huber, B. Tarasuik, J., Antoniou, M. N., Garret, C., Bowe, S. J., Kaufman, J., & The Swinburne Babylab Team. (2016). Young children's transfer of learning from a touchscreen device.

Computers in Human Behavior, 56, 56-64. DOI: <https://doi.org/10.1016/j.chb.2015.11.010>

Hughes, C. H., & Ensor, R. A. (2009). How do families help or hinder the emergence of early executive function? In C. Lewis & J. I. M. Carpendale (Eds.), *Social interaction and the development of executive function. New Directions in Child and Adolescent Development*, 123, 35–50.

Kankaanranta, M., Koivula, M., Laakso, M., & Mustola, M. (2017). Digital games in early childhood: Broadening definitions of learning, literacy, and play. In *Serious Games and Edutainment Applications*, 349-367, Springer, Cham

Kannass K.N. & Colombo, J. (2007) The Effects of Continuous and Intermittent Distractors on Cognitive Performance and Attention in Preschoolers, *Journal of Cognition and Development*, 8:1, 63-77. DOI: <http://dx.doi.org/10.1080/15248370709336993>

Klahr, D., Triona, L. M., & Williams, C. (2007). Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science Teaching*, 44(1), 183-203. DOI: 10.1002/tea.20152

Kloos, H. & Van Orden, G. C. (2005). Can a preschooler's mistaken belief benefit learning? *Swiss Journal of Psychology*, 64(3), 195-205. DOI: 10.1024/1421-0185.64.3.195

Lieberman, D., Bates, C., & So, J. (2009). Young children's learning with digital media, computers in the schools. *Computers in the schools*, 26(4), 271-283.

DOI:10.1080/07380560903360194

- Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology*, 36(5), 596-613. DOI: 10.1037//OOI2-1649.36.5.596
- Mahy, C. E. V., Moses, L. J., & Kliegel, M. (2014). The impact of age, ongoing task difficulty, and cue salience on preschoolers' prospective memory performance: The role of executive function. *Journal of Experimental Child Psychology*, 127, 52-64. DOI: <https://doi.org/10.1016/j.jecp.2014.01.006>
- Marcovitch, S., Boseovski, J. J., & Knapp, R. J. (2007). Use it or lost it: Examining preschoolers' difficulty in maintaining and executing a goal. *Developmental Science*, 10(5), 599-564. DOI: 10.1111/j.1467-7687.2007.00611.x
- Masterman, E. & Rogers, Y. (2002). A framework for designing interactive multimedia to scaffold young children's understanding of historical chronology, *Instructional Science*, 30(3), 221-241. DOI: 10.1023/A:1015133106888
- Molnar, M. & Cavanagh, S. (2013). Consumer demand for digital learning games, simulations growing worldwide: But level of interest in K-12 lags behind, Education Week. Retrieved from <http://www.edweek.org/ew/articles/2013/09/18/04games.h33.html>
- Nikoyalev, M., Reich, S. M., Musket, T., Tadjbakhsh, N., Callaghan, M. N. (under review). The use of feedback in online "educational" games for preschoolers.
- Obikwelu, C., Read, J., & Sin, G. (2012). The scaffolding mechanism in serious games. *Fun and Game*, 9, 4-6. ISBN: 978-2- 917490-21- 1
- Park, S. & Kim, C. (2015). Boosting learning-by-teaching in virtual tutoring. *Computers in Education*, 82, 129-140. <http://dx.doi.org/10.1016/j.compedu.2014.11.006>

- Petersen, L. A., & McNeil, N. M. (2013). Effects of perceptually rich manipulatives on preschoolers' counting performance: Established knowledge counts. *Child Development*, 84(3), 1020-1033. DOI: 10.1111/cdev.12028
- Pew Research Center. (2017). A third of Americans live in a household with three or more smartphones. Retrieved from: <http://www.pewresearch.org/fact-tank/2017/05/25/a-third-of-americans-live-in-a-household-with-three-or-more-smartphones/>
- Piaget, J. (1984). The Attainment of Invariants and Reversible Operations in the Development of Thinking. *Social Research*, 51(1/2), 167-184. Retrieved from: <http://www.jstor.org/stable/40970935>
- Pianta, R. C., Barnett, W. S., Burchinal, M., & Thornburg, K. R. (2009). The effects of preschool education: What we know, how public policy is or is not aligned with the evidence base, and what we need to know. *Psychological Science in the Public Interest*, 10(2), 49-88. DOI: 10.1177/1529100610381908
- Plowman, L., McPake, J., & Stephen, C. (2010). The technologisation of childhood? Young children and technology in the home. *Children and Society*, 24(1), 63-74. DOI: 10.1111/j.1099-0860.2008.00180.x
- Reich, S.M. & Black, R. (2012). Missed opportunities on Webkinz when developmental abilities are considered. *Journal of Applied Developmental Psychology*, 33, 136-145. DOI: 10.1016/j.appdev.2012.02.001
- Romeo, G., Edwards, S., McNamara, S., Walker, I., & Ziguras, C. (2003). Touching the screen: Issues related to the use of touchscreen technology in early childhood education. *British Journal of Educational Technology*, 34(3), 329-339. DOI: 10.1111/1467-8535.00330
- Rutherford, T., Farkas, G., Duncan, G., Burchinal, M., Kibrick, M., Gaham, J., ... Martinez, M.

- (2014). A randomized trial of an elementary school mathematics software intervention: Spatial-temporal math. *Journal of Research on Educational Effectiveness*, 7, 358-383. DOI: 10.1080/19345747.2013.856978
- Sarnecka, B. (2015). Learning to represent exact numbers. Synthese, *An International Journal for Epistemology, Methodology, and Philosophy of Science*. DOI: 10.1007/s11229-015-0854-6
- Sharma, P. & Hannafin, M., (2007). Scaffolding in technology-enhanced learning environments. *Interactive Learning Environments*, 15(1), 27-46. DOI: 10.1080/10494820600996972
- Shoukry, L., Sturm, C., & Galal-Edeen, G.H. (2015). Pre-MEGa: a proposed framework for the design and evaluation of preschoolers' mobile educational games. In *Innovations and Advances in Computing, Informatics, Systems Sciences, Networking and Engineering*, 385-390. Springer, Cham. DOI: 10.1007/978-3-319-06773-5_52
- Shuler, C. (2012). iLearn II: An addendum—An analysis of the games category of Apple's App Store. Retrieved from: http://www.joanganzcooneycenter.org/wp-content/uploads/2012/09/jgcc_ilearn_games_addendum.pdf
- Thorell, L., Lindqvist, S., Nutley, S., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, 12(1), 106-113. DOI: 10.1111/j.1467-7687.2008.00745.x
- Vaala, S., Ly, A., & Levine, M.H. (2015). Getting a read on the app stores: A market scan and analysis of children's literacy apps. New York, NY: The Joan Ganz Cooney Center at Sesame Workshop.
- Vatavu, R., Cramariuc, G. & Schipor, D.M. (2015). Touch interaction for children 3 to 6years: Experimental findings and relationship to motor skills. *International Journal of Human-*

Computer Studies, 74, 54-76. DOI: <http://dx.doi.org/10.1016/j.ijhcs.2014.10.007>

Vygotsky, L. (1978). Interaction between learning and development. From: *Mind and Society* (79-91). Cambridge, MA: Harvard University Press

Ward, W., Cole, R., Bolaños, D., Buchenroth-Martin, C., Svirsky, E., & Weston, T. (2013). My Science Tutor: A Conversational Multimedia Virtual Tutor. *Journal of Educational Psychology*. Advance online publication. 105(4), 1115. DOI: 10.1037/a0031589

Wouters, P. & van Oostendorp, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers & Education*, 60, 412-425. DOI: <http://dx.doi.org/10.1016/j.compedu.2012.07.018>

Zichermann, G. & Cunningham, C. (2011). *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*. Sebastopol, CA: O & Reilly Media, Inc.

APPENDIX A

Study 1 Coding Rubric

Q1 Coder Name

- ☐ Melissa (1)
- ☐ Amrit (2)
- ☐ Annie (3)
- ☐ Jannell (4)
- ☐ Other (specify) (5) _____

Q2 Game Name

Q3 Platform

- ☐ Apple tablet (1)
- ☐ Android tablet (2)
- ☐ Other: (3) _____

Q4 Was this app free or purchased?

- ☐ Free (1)
- ☐ Purchased (2)
- ☐ Other (specify): (3) _____

Q5 App's Target Content

- ☐ This is a math app (1)
- ☐ This is a literacy app (2)
- ☐ This app is a mixture of math and literacy (3)
- ☐ This app teaches math and additional subject(s) (4)
- ☐ This app teaches literacy and additional subject(s) (5)
- ☐ This app teaches math, literacy, and additional subject(s) (6)

Q6 Menu options

- ☐ There is no menu, it just jumps into play. (1)
- ☐ There is a menu, with a single play button (2)
- ☐ There is a menu, with multiple button options (3)

Q7 Game plot

- ☐ Non-integrative plot: There is a plot presented before starting the puzzle(s), but the player is not integrated into the storyline. (1)
- ☐ Integrative plot: The game builds a story before starting the puzzle(s) and integrates the player's role into the storyline. (2)
- ☐ There is no plot (3)

Q8 Description: Simple description of what you see when the game starts to play (e.g., group of friends sitting at picnic table.)

Q9 Applicability

- ☐ Free Play (Not Applicable): There is no right or wrong answer. (1)
- ☐ Assesses Input (Applicable): There is a right or wrong answer and the game responds to inputted answers. (2)

Q10 Single Puzzle Type

- ☐ Not present (1)
- ☐ One game only. No variation to progress further. (2)
- ☐ Predetermined pathway of progression. (3)
- ☐ Selection choice in pathway progression. (4)
- ☐ Mixture of predetermined pathway and choice of progression (5)

Q11 Multiple Puzzle Type

- ☐ Not present (1)
- ☐ Pre-selected order of play (e.g., the game just starts playing and leads to a sequence of different puzzles) (2)
- ☐ Player had choice of play (e.g., an opening menu of puzzle options to select from) (3)
- ☐ Mixture of pre-selected and player chosen play (4)

Q12 Leveling

- ☐ Not present (1)
- ☐ Subtle levels of play that increase in challenge (e.g., no specific level 1 or 2 but gradually gets harder) (2)
- ☐ Obvious levels of play that increase in challenge (e.g., number of points/puzzles leads to increased level number) (3)
- ☐ Subtle and obvious levels (4)
- ☐ Other (specify): (5) _____

Q13 Level of difficulty can be selected (e.g., easy, medium, and hard)

- ☐ Yes (1)
- ☐ No (2)

Q14 How many puzzles are in the game?

- ☐ 1 (1)
- ☐ 2 (2)
- ☐ 3 (3)
- ☐ 4 (4)
- ☐ 5 (5)
- ☐ 6 (6)
- ☐ 7 (7)
- ☐ 8 (8)
- ☐ 9 (9)
- ☐ 10 (10)

Q15 Puzzle ONE: Puzzle 1 Premise: Does the PUZZLE provide some sort of storyline to make the game meaningfully connect to the player? (Add description of story in Additional Notes)

- ☐ No plot: There is no storyline. It jumps right into play. (1)
- ☐ Non-integrative plot: There is a plot presented at the start of the puzzle, but the player is not integrated into the storyline. (2)
- ☐ Integrative plot: The puzzle builds a story and integrates the player's part into the storyline. (3)
- ☐ Other (specify) (4) _____

Q16 Puzzle 1: Oral/Textual instructions at the start of a task/puzzle

- ☐ No instructions/Preparation ("Let's go!") (1)
- ☐ Minimal instruction ("How many chickens do you see?") (2)
- ☐ Moderate instruction ("Give each friend at the table a cookie.") (3)
- ☐ Elaborate instruction ("Drag the correct number of cookies to each player.") (4)
- ☐ Other (specify): (5) _____

Q17 Puzzle 1: Guidance for completing task

- ☐ No guidance is provided for user to know how to complete a task. User is expected to figure it out. (1)
- ☐ The app gives instructions for completing task ("Click the green button when you're done.") (2)
- ☐ No guidance is needed. It immediately indicates if it's right or wrong. (3)

Q18 Puzzle 1: Timer present?

- ☐ No (1)
- ☐ Yes (2)
- ☐ Other (specify): (3) _____

Q19 Puzzle 1: Format of Instructions

- ☐ Textual (1)
- ☐ Oral (2)
- ☐ Visual (visually things change or are highlighted while oral instructions are given) (3)
- ☐ Other (specify) (4) _____

Q20 Puzzle 1: Instructions in the middle of a task

- ☐ No automatic instructions given after 1 minute of no participation (1)
- ☐ Automatic repeated verbatim instructions after up to 1 minutes of no participation (2)
- ☐ Automatic rephrased instructions after up to 1 minutes of no participation (3)
- ☐ Other (specify): (4) _____

Q21 Puzzle 1: Characters

	No (1)	Yes (2)
Play in 1st person view (Q16_1)	☐	☐
Controlling play as Avatar (Q16_2)	☐	☐
Avatar present - doesn't responds to actions (Q16_3)	☐	☐
Avatar present AND responds to actions (Q16_4)	☐	☐
Controlling play as pre-made character (Q16_6)	☐	☐
Pre-made character present - doesn't respond to actions (Q16_5)	☐	☐
Pre-made character present AND responds to actions (Q16_7)	☐	☐
At least 1 NPC on screen during play (Q16_9)	☐	☐
Other (Q16_8)	☐	☐

Q22 Puzzle 1: Character Modeling

	No (1)	Yes (2)
NPC pre-models correct (Q21_1)	☐	☐
NPC pre-models incorrect (Q21_2)	☐	☐

Q23 Puzzle 1: Play Modeling

- ☐ No play modeling (1)
- ☐ Shows a hand/mouse/keyboard to demonstrate how to play (before play) (2)
- ☐ Has an arrow or hand hinting which direction to click next (during play) (3)
- ☐ Other (specify) (4) _____

Q24 Puzzle 1: Teaches math/literacy content outside of "play-time" (Basically, they demonstrate something, such as counting the number of coins collected, which is educational, but not part of the tasks players complete during "play")

- ☐ No (1)
- ☐ Yes (2)
- ☐ Other: (3) _____

Q25 Puzzle 1: Primary MATH Learning Goals: The goals of the game explicitly intended by designers

- ☐ N/A - It's a literacy game (1)
- ☐ Number Tracing: Outline the number symbols given (2)
- ☐ Number Symbol Recognition: Identify the names of number symbols displayed (3)
- ☐ Cardinality (knowing how many numbers are in a set or objects are in a group; e.g., "2, 4, 6" has 3 numbers) (4)
- ☐ Counting numbers/objects (5)
- ☐ Using a number-line (6)
- ☐ Adding Numbers: Number symbols added together (7)
- ☐ Adding Objects: Objects added together (8)
- ☐ Subtracting Numbers: Number symbols subtracted from each other (9)
- ☐ Subtracting Objects: Objects subtracted from each other (10)
- ☐ Patterns of Numbers: Sort or identify patterns of number symbols/quantities (sequence them in order) (11)
- ☐ Patterns of Objects: Sort or identify patterns of objects (12)
- ☐ Patterns of Colors: Sort or identify patterns of colors (13)
- ☐ Patterns of Shapes: Sort or identify patterns of shapes (14)
- ☐ Identify Shapes: Indicating which is a triangle, square, circle, etc. (15)
- ☐ Identify Physical Measurable Attributes: Identify sizes of objects/shapes (16)
- ☐ Compare Numerical Measurable Attributes: Compare sizes of amounts (e.g., "Which pen has less chicks?") (17)
- ☐ Identifying even or odd numbers (doesn't just tell you they're even or odd, but actually requires you to choose) (18)
- ☐ Currency (Counting, adding, subtracting, etc.) (19)
- ☐ Measurement (20)
- ☐ Other (specify) (21) _____

Q26 Puzzle 1: Secondary MATH Learning Goals: The goals of the game implicitly intended by designers

- ☐ N/A - It's a literacy game (1)
- ☐ No Secondary Learning Goals in this puzzle (0)
- ☐ Number Tracing: Outline the number symbols given (2)
- ☐ Number Symbol Recognition: Identify the names of number symbols displayed (3)
- ☐ Cardinality (knowing how many numbers are in a set or objects are in a group; e.g., "2, 4, 6" has 3 numbers) (4)
- ☐ Counting numbers/objects (5)
- ☐ Using a number-line (6)
- ☐ Adding Numbers: Number symbols added together (7)
- ☐ Adding Objects: Objects added together (8)
- ☐ Subtracting Numbers: Number symbols subtracted from each other (9)
- ☐ Subtracting Objects: Objects subtracted from each other (10)
- ☐ Patterns of Numbers: Sort or identify patterns of number symbols/quantities (sequence them in order) (11)
- ☐ Patterns of Objects: Sort or identify patterns of objects (12)
- ☐ Patterns of Colors: Sort or identify patterns of colors (13)
- ☐ Patterns of Shapes: Sort or identify patterns of shapes (14)
- ☐ Identify Shapes: Indicating which is a triangle, square, circle, etc. (15)
- ☐ Identify Physical Measurable Attributes: Identify sizes of objects/shapes (16)
- ☐ Compare Numerical Measurable Attributes: Compare sizes of amounts (e.g., "Which pen has less chicks?") (17)
- ☐ Identifying even or odd numbers (doesn't just tell you they're even or odd, but actually requires you to choose) (18)
- ☐ Currency (Counting, adding, subtracting, etc.) (19)
- ☐ Measurement (20)
- ☐ Other (specify) (21) _____

Q27 Puzzle 1: Primary LITERACY Learning Goals: The goals of the game explicitly intended by designers

- ☐ N/A - It's a math game (1)
- ☐ Alphabetical order (e.g., Put letters in alphabetical order) (2)
- ☐ Trace Letters (e.g., An outline of a letter is on the screen and user follows the lines) (3)
- ☐ Letter writing (e.g., Write letters without an image guiding user to trace) (4)
- ☐ Identify letters (e.g., Find the letter A) (5)
- ☐ Uppercase vs. Lowercase (e.g., Make upper vs. lowercase distinction) (6)
- ☐ Story comprehension (e.g., Put pieces of the storybook back into the order it was read) (7)
- ☐ Text/word comprehension (e.g., Complete the words with these letter combinations) (8)
- ☐ Picture naming (e.g., Tap on the cat, tap on the objects with spots, tap on opposite picture) (9)
- ☐ Phonological awareness of sounds (e.g., Sound out letters to add a letter in to make a word) (10)
- ☐ Alphabet knowledge (e.g., Find the words/images that begin with letter "C") (11)
- ☐ Oral language comprehension (e.g., Identify text with word said) (12)
- ☐ Unscramble letters to make a word (13)
- ☐ Complete the sentence (e.g., Fill in missing words) (14)
- ☐ Differentiate vowels and consonants (15)
- ☐ Other (specify): (16) _____

Q28 Puzzle 1: Secondary LITERACY Learning Goals: The goals of the game implicitly intended by designers

- ☐ N/A - It's a math game (1)
- ☐ No secondary learning goals (0)
- ☐ Alphabetical order (e.g., Put letters in alphabetical order) (2)
- ☐ Trace Letters (e.g., An outline of a letter is on the screen and user follows the lines) (3)
- ☐ Letter writing (e.g., Write letters without an image guiding user to trace) (4)
- ☐ Identify letters (e.g., Find the letter A) (5)
- ☐ Uppercase vs. Lowercase (e.g., Make upper vs. lowercase distinction) (6)
- ☐ Story comprehension (e.g., Put pieces of the storybook back into the order it was read) (7)
- ☐ Text/word comprehension (e.g., Complete the words with these letter combinations) (8)
- ☐ Picture naming (e.g., Tap on the cat, tap on the objects with spots, tap on opposite picture) (9)
- ☐ Phonological awareness of sounds (e.g., Sound out letters to add a letter in to make a word) (10)
- ☐ Alphabet knowledge (e.g., Find the words/images that begin with letter "C") (11)
- ☐ Oral language comprehension (e.g., Identify text with word said) (12)
- ☐ Unscramble letters to make a word (13)
- ☐ Complete the sentence (e.g., Fill in missing words) (14)
- ☐ Differentiate vowels and consonants (15)
- ☐ Other (specify): (16) _____

Q29 Puzzle 1: Tertiary learning Goals (Not Math/Literacy related)

- ☐ No Tertiary Learning Goals in this puzzle (1)
- ☐ Sharing (2)
- ☐ Helping others in need (e.g., "Oh no, we dropped our coins! Can you help us find them?") (3)
- ☐ Waiting patiently (4)
- ☐ Hand-eye coordination (5)
- ☐ Consumerism/purchasing (6)
- ☐ Pretend school (7)
- ☐ Persistence (e.g., "you can always figure it out if you keep trying") (8)
- ☐ Inhibitory control (e.g., All shapes are falling and you have to ignore the others and only identify the circles) (9)
- ☐ Personalizing (e.g., avatar or decorate your own train) (10)
- ☐ Other (specify) (11) _____

Q30 Puzzle 1: Format of math/literacy-related questions/tasks

- ☐ Free write/Open ended (1)
- ☐ Multiple choice options to choose from (2)
- ☐ Mixture of free-write and multiple choice (3)

Q31 ALL CORRECT PLAY: Only enter correct answers in the game to complete this section.

Puzzle 1: Sound Feedback (CORRECT)

- ☐ No sound effect if an answer is correct (1)
- ☐ Small sound (e.g., bell) indicates answer is correct (2)
- ☐ Positive emotional sound (e.g., giggle) (3)
- ☐ Moderate sound (e.g., celebratory music) indicates answer is correct (4)
- ☐ Other (specify) (5) _____

Q32 Puzzle 1: Visual Feedback (CORRECT)

- ☐ Visually, nothing changes. Immediately switches to next task/puzzle. (1)
- ☐ Visually demonstrates answer was right (e.g., circle around selection or highlights selected answer in green) (2)
- ☐ Positive visual emotion/praise (e.g., balloons fly, characters smile, stars) (3)
- ☐ Neutral visual changes that are not scaffolding (4)
- ☐ General guiding hints that lead player towards the correct answer (e.g., dotted lines to friends who still need a cupcake, crosses out wrong answers and leaving fewer possible choices, etc.) (5)
- ☐ Scaffolded visual feedback (e.g., oranges highlight one by one, to count total number of oranges) (6)
- ☐ Other (specify) (7) _____

Q33 Puzzle 1: Oral Feedback (CORRECT)

- ☐ No oral feedback for correct (1)
- ☐ Positive oral praise (e.g., "Great job!") (2)
- ☐ Narrates action (e.g., "That's right! You selected 3 oranges!") (3)
- ☐ Scaffolded oral feedback (e.g., "Correct! One...two...three...that's three oranges!") (4)
- ☐ Unrelated oral elaboration (e.g., "That's right! And square starts with the letter S.") (5)
- ☐ Other (specify) (6) _____

Q34 Puzzle 1: Textual Feedback (CORRECT)

- ☐ No textual feedback for correct (1)
- ☐ Positive textual praise ("Yay! You did it!") (2)
- ☐ Scaffolded textual feedback ("One...two...three. Correct! There are three toy trucks.") (3)
- ☐ Unrelated textual elaboration ("Boats have triangles too.") (4)
- ☐ Other (specify) (5) _____

Q35 Puzzle 1: Point System (CORRECT)

- ☐ No point system (1)
- ☐ High score achievement. The number of points received is the reward. (2)
- ☐ Receives numerical points for correct that are put towards something like leveling up the character (3)
- ☐ Collectible items (doesn't count towards a score) (4)
- ☐ Collectible items (leads to a reward, increased score, or leveling up) (5)
- ☐ Collects prizes/rewards at the end of tasks (e.g., stickers for a sticker book) (6)
- ☐ Items or points unlock prizes (7)
- ☐ Items or points unlock levels (8)
- ☐ Other (specify) (9) _____

Q36 ALL WRONG PLAY: Only enter wrong answers in the game to complete this section Puzzle 1: Sound Feedback (WRONG)

- ☐ No sound effect if answer is wrong (1)
- ☐ Small sound (e.g., a horn, beep, mm-m) indicates answer is wrong (2)
- ☐ Moderate negative sound (e.g., sad music, crying) indicates answer is wrong (3)
- ☐ Other (specify) (4) _____

Q37 Puzzle 1: Visual Feedback (WRONG)

- ☐ Visually, nothing changes when the player gets the answer wrong. Task just switches. (1)
- ☐ Neutral visual feedback (e.g., characters fidget around after answer wrong) (2)
- ☐ Visually demonstrates answer was wrong (e.g., X over selection or highlights selected answer in red) (3)
- ☐ Demonstrates correct answer, then moves on (4)
- ☐ Subtle negative emotional feedback (e.g., sad faces) (5)
- ☐ Guiding visual hints to get player closer to correct answer (e.g., dotted lines leading to the number of friends left without a cupcake) (6)
- ☐ Highlights correct answer (e.g., flashes correct answer to get you to select it) (7)
- ☐ Other (specify) (8) _____

Q38 Puzzle 1: Oral Feedback (WRONG)

- ☐ No oral feedback for wrong (1)
- ☐ Non-scaffolded oral feedback ("Not quite. Try again.") (2)
- ☐ Gives correct answer (3)
- ☐ Repeated verbatim instruction ("No, we need 3 cookies. Can you count 3 cookies?") (4)
- ☐ Rephrased instructions ("No, that's a square. Can you find the circle? We need a round toy.") (5)
- ☐ Scaffolded oral feedback ("Not quite. A circle has no angles. The square you selected has 4 angles.") (6)
- ☐ Unrelated oral elaboration ("Find the square. Square starts with the letter 's'.") (7)
- ☐ Other (specify) (8) _____

Q39 Puzzle 1: Textual Feedback (WRONG)

- ☐ No textual feedback for wrong (1)
- ☐ Non-scaffolded textual feedback ("Nope.") (2)
- ☐ Scaffolded textual feedback ("One...two...three. There are only 3 trucks. We need 4.") (3)
- ☐ Unrelated textual feedback ("Boats have triangles.") (4)
- ☐ Repeated verbatim instruction ("No, we need 3 cookies. Can you count 3 cookies?") (5)
- ☐ Rephrased instructions ("No, that's a square. Can you find the circle? We need a round toy.") (6)
- ☐ Other (specify) (7) _____

Q40 Puzzle 1: Indicates that the player was wrong.

- ☐ Does not visually, verbally, or orally indicate that the player got it wrong ("Let's see what happens when we select this one instead...This is the correct answer!") (1)
- ☐ Indicates that the player was wrong ("Uh oh. That's not right") (2)
- ☐ Other (specify) (3) _____

Q41 Puzzle 1: Point System (WRONG)

- ☐ Nothing happens to points after wrong answers (1)
- ☐ Loses points for getting answers wrong (2)
- ☐ Receives points for effort (even if wrong) that become part of an overall score (3)
- ☐ Receives points for effort (even if wrong) that become part of leveling up (4)
- ☐ Receives prize for effort (right or wrong answers are irrelevant in this case) (5)
- ☐ Missed collectible items do not deduct points (6)
- ☐ Missed collectible items do deduct points (7)
- ☐ Other (specify) (8) _____

Q42 Puzzle 1: Attempts at answering

- ☐ Unlimited: Answer until correct - Can get wrong answer forever and never move forward (1)
- ☐ One attempt only. (e.g., Count number of dinosaurs you see. You count 4. They show that there are actually 3 dinosaurs. Then it moves to next task) (2)
- ☐ Limited incorrect attempts: Switches to answer and/or different task after entering so many wrong answers (3)
- ☐ Other (specify) (7) _____

Q43 Puzzle 1: Doesn't allow you to over-estimate. Once you reach the right number/amount, it says you're correct

- ☐ True (1)
- ☐ False (2)
- ☐ N/A (e.g., it's a literacy app or it doesn't come up) (3)

Q44 Puzzle 1: Scaffolded challenge

- ☐ There was no scaffolded challenge in this puzzle (1)
- ☐ Challenge decreased within puzzle (e.g., adding big numbers, but after getting many wrong, the numbers get smaller to make it easier to add) (2)
- ☐ Challenge increased within puzzle (e.g., after getting addition problems right, the numbers get bigger) (3)
- ☐ Challenge decreased and increased (4)
- ☐ Other (specify) (5) _____

Q45 Is there another puzzle to code?

- ☐ No (1)
- ☐ Yes (name) (2) _____

... repeats for up to 10 puzzles.

Q324 Whole Game: Scaffolded challenge

- ☐ Level of challenge did not appear to increase or decrease in challenge ever (1)
- ☐ Challenge increased across puzzles (e.g., the first puzzle was counting, the next was addition) (2)
- ☐ Challenge decreased across puzzles (e.g., after getting number recognition problems wrong, it takes you back to tracing numbers) (3)
- ☐ There was an increase in non-math related challenge of the game (e.g., more difficult to collect items/points) (4)
- ☐ Other (specify) (5) _____

Q325 Overall Reward for completion

- ☐ No (no reward) (1)
- ☐ Yes (e.g., there's a coloring activity at the end of the game; buy a bear at the end) (2)
- ☐ Other (3) _____

Q326 Was the app multi-player?

- ☐ No. Individual play only. (1)
- ☐ Yes, you play against NPCs in the game (2)
- ☐ Yes, you can play against friends in the same room (3)
- ☐ Yes, you can play as a team with friends in the same room (4)
- ☐ Yes, you can play against friends online (5)
- ☐ Yes, you can play as a team with friends online (6)
- ☐ Other (specify): (7) _____

Q327 Utilizing Special Tablet Function

	No (1)	Yes (2)
Touch screen dragging (Q32_1)	☐	☐
Touch screen tracing (Q32_2)	☐	☐
Multitouch (Q32_3)	☐	☐
Zooming in and out (Q32_4)	☐	☐
Uses accelerometer of tablet (Q32_5)	☐	☐
Camera access (Q32_6)	☐	☐
Voice recording (Q32_7)	☐	☐

Q328 Any additional tablet functions used?

Q329 Were there in-app purchases offered for additional puzzles in this game?

- ☐ No (1)
- ☐ Yes (2)
- ☐ Other (specify): (3) _____

Q330 Were there in-app purchases offered for additional in-game features?

- ☐ No (1)
- ☐ Yes (2)
- ☐ Other (specify): (3) _____

Q331 Enjoyment Measure

	No (1)	Yes (2)
The overarching goal of the game is initially presented	☐	☐
The goals of the puzzle(s) are made clear throughout the game	☐	☐
There are multiple goals in each puzzle	☐	☐
There are multiple goals in game overall (e.g., puzzle 1 counting, puzzle 2 addition)	☐	☐
The game provides opportunity for user to control game play (select levels/tasks)	☐	☐
The player is quickly and easily involved in game play	☐	☐
Content is embedded in fantasy contexts - Talking animals	☐	☐
Content is embedded in fantasy contexts - Mystical creatures	☐	☐
Content is embedded in fantasy contexts - Media-based characters/settings	☐	☐
The game includes unexpected and interesting things for the user	☐	☐
The game uses humor at any point in the game (prompt, feedback, context, etc.)	☐	☐
Encourages imagination to encourage identification with characters or contexts.	☐	☐
Allows personalization of an “avatar”	☐	☐
The game includes sound and visual effects	☐	☐
Variable auditory and visual effects are used as feedback	☐	☐
Every puzzle relates to a single theme	☐	☐
The game applies pressure to the player by using a timer	☐	☐
The user gets frequent feedback	☐	☐
There are multiple optimal winning strategies	☐	☐
The game gives rewards (points, prizes, leveling-up character, etc.)	☐	☐
There’s an exit button to easily exit puzzles (e.g., large red “X” or menu button)	☐	☐
The puzzle(s) are/is replayable	☐	☐

Q332 Final Notes (optional)

APPENDIX B

Tables of Study 1 Results

Below are tables representing the reported frequencies of app features used in the popular preschool apps examined in Chapter 2. Since it was possible for individual apps to include multiple features from the same category, all reported frequencies are mutually exclusive, so percentages do not add up to 100%.

Simplicity of Content and Clarity of Goals

Learning Goals and Features of Enjoyment	
Clear Goals	79.5%
Single Theme	87.1%
Math Apps Learning Goals	
Most Common Math App Learning Goals	
Counting	50.5%
Number Recognition	40.5%
Least Common Math App Learning Goals	
Even/Odd Numbers	0.01%
Currency	0.01%
Other Math App Learning Goals	
Number Tracing	12%
Adding Numbers	17%
Adding Objects	13%
Subtracting Numbers	11%
Subtracting Objects	7%
Identify Shapes	23%
Patterns of Numbers	15%
Literacy Apps Learning Goals	
Most Common Literacy App Learning Goals	
Identifying Letters	33.0%
Phonological Awareness	26.6%
Least Common Literacy App Learning Goals	

Vowels/Consonants	0.2%
Story Comprehension	0.01%
Completing Sentences	0.01%
Other Literacy App Learning Goals	
Alphabetical Order	9%
Trace Letters	16%
Picture Naming	8%
Letter Writing	2%
Text/Word Comprehension	11%

Types of Prompts	
No Instructions Provided	18.7%
Moderate Instructions (e.g., How many apples are in the basket?)	32.8%
Elaborate Instructions (e.g., How many apples are in the basket? Tap on the correct number.)	25.5%
Verbal Prompts	79%
Verbal with Text Prompts	35.7%
Prompts with Visual Cues	13.5%

Modeling Goals	
Character Modeling	4.7%
Play Demonstration (e.g., hand or arrow)	17%
Next Step Hints (e.g., When you're finished, tap on the green button.)	10.5%

In-Play Guidance After Long Pause	
Repeated Instructions	15.3%
Rephrased Instructions	6.4%

Quality of Feedback and Rewards

Types of Feedback
Correct Answers

Sound-based Feedback	28.7%
Positive Sound Effects	28.1%
Positive Verbal Feedback	73.7%
No Text Feedback	84.2%
Positive Text Feedback	23.4%
Visual Feedback	73.1%
Celebratory Visual Feedback	68.4%
Incorrect Answers	
Sound Effects	65.5%
Non-scaffolded Verbal Feedback	46.8%
No Text Feedback	94.5%
Visual Indication of Error (e.g., A red “X” over selected answer.)	57.3%
Neutral Visual Feedback (e.g., Character stands in new position.)	35.7%

Consistency of Guidance	
Pre-play and Post-response Guidance	8.2%

Rewards System	
Educational Rewards	3.5%
General Rewards	49.7%
Prizes (e.g., stickers and badges)	31%
Leveling up Characters	1.8%
Unlocking Levels	1.8%

Structure of Challenge

Leveling	
No Leveling	72.5%
Subtle Leveling	15.2%
Obvious Leveling	9.9%
Both Subtle and Obvious Leveling	2.3%

Scaffolded Challenge	
-----------------------------	--

No Scaffolded Challenge Within Puzzles	68.4%
Increased Challenge Within Puzzles	30%
No Adaptive Leveling	80.7%
Increased Leveling When Correct	17%
Decreased Leveling When Incorrect	0%

Additional Mobile App-based Interactions

Touchscreen Interactions Used	
Dragging	65%
Tracing	27%
Multi-touch	2.3%
Pinching Zoom	0%

Mobile Device Capabilities	
Accelerometer	3%
Camera	1.2%
Voice Recorder	2.3%

APPENDIX C

Participants' Background Questionnaire

1. Child Identification:
 - a. Child's Date of Birth (mm/dd/yyyy) _____ / _____ / _____ Age: _____
 - b. Child is: ☐ Female ☐ Male
2. What is your child's race? (check all that apply):
 - ☐ American Indian or Alaskan Native
 - ☐ Asian or Asian American
 - ☐ Black or African American
 - ☐ Hispanic or Latino
 - ☐ Native Hawaiian or other Pacific Islander
 - ☐ White or Caucasian
 - ☐ Other - please specify _____
 - ☐ Decline to state
3. Is English the main language used in your household?
 - ☐ No - the main language(s) is/are: _____
 - ☐ Yes, only English
 - ☐ Yes, and _____
4. How often does your child use a tablet (e.g., iPad, Kindle)?
 - ☐ Never
 - ☐ Less than once a week
 - ☐ A couple days of the week (1-2 days)
 - ☐ A few times a week (3-4 days)
 - ☐ Most days of the week (5-6 days)
 - ☐ Every day of the week (7 days)
5. Does your child ever play educational games on a tablet?
 - ☐ No (skip to question 6)
 - ☐ Yes
 - a. How often does your child play educational games on a tablet?
 - ☐ Less than once a week
 - ☐ A couple days of the week (1-2 days)
 - ☐ A few times a week (3-4 days)
 - ☐ Most days of the week (5-6 days)
 - ☐ Every day of the week (7 days)

b. If yes, does your child typically play with (check all that apply):

- ☐ A sibling
- ☐ A relative
- ☐ A friend
- ☐ A parent/guardian
- ☐ Independently

c. If yes, what type of educational games does your child play on a tablet? (check all that apply):

- ☐ Math (e.g., counting, adding, measurements)
- ☐ Spatial (e.g., shape recognition, jigsaw puzzles)
- ☐ Color learning
- ☐ Literacy/Reading
- ☐ Science (e.g., coding, physics ideas)
- ☐ Social skills (e.g., sharing, taking turns)
- ☐ Other (please specify): _____

6. What is your current marital status?

- ☐ Single or never married
- ☐ Married/living as married
- ☐ Separated
- ☐ Divorced
- ☐ Widowed

7. What is your highest level of education?

- ☐ No high school
- ☐ Some high school
- ☐ High school graduate
- ☐ Some college, but did not graduate
- ☐ Two-year college degree or technical certificate
- ☐ Four-year college degree
- ☐ Some graduate/professional school
- ☐ Graduate or professional school degree

8. Please mark the estimated yearly household income:

- ☐ \$0 to \$24,999
- ☐ \$25,000 to \$49,999
- ☐ \$50,000 to \$74,999
- ☐ \$75,000 to \$124,999
- ☐ \$125,000 to \$199,999
- ☐ \$200,000 +

APPENDIX D

Study 2 Game Details

Feedback Tasks

1. Plickatoo 
 - a. Prompts:
 - i. “First, we need to round up all the Plickatoos. Plickatoos are aliens that are green with spots. Move all of the plickatoos to their spaceship!”
 - ii. “See any more Plicktoos? Put them on the spaceship!”
 - iii. “A few more Plickatoos are hanging out here. Can you take them to the spaceship?”
 - b. Scaffolded Feedback:
 - i. Correct: “That’s a plickatoo - it’s green!”
 - ii. Incorrect: “That’s not green”; “Plickatoos are green”
2. Gonk 
 - a. Prompts:
 - i. “Next it’s the Gonks turn to board. Move all the Gonks to their spaceship!”
 - ii. “Here are some more Gonks. Drag them to the spaceship!”
 - iii. “Do you see any more Gonks? Put all the Gonks on the spaceship!”
 - b. Scaffolded Feedback:
 - i. Correct: “You found the blue Gonk!”
 - ii. Incorrect: “Remember, a Gonk is square”; “We need to find the blue Gonks.”
3. Flaroo 
 - a. Prompts:
 - i. “Here are more aliens. Drag all of the Flaroos to the spaceship!”
 - ii. “Still some more Flaroos over here. Put the Flaroos on the spaceship!”
 - iii. “Do you see anymore Flaroos? Drag them to the spaceship!”
 - b. Scaffolded Feedback:
 - i. Correct: “That Flaroo is bright red!”
 - ii. Incorrect: “Flaroos have black ears.”; “Remember, a Flaroo is bright red.”

Leveling Tasks:

1. Froodle

a. Prompts:

- i. "This is a Froodle. Froodles are orange and shaped like a square. That means they have four corners. Move the Froodle to the spaceship!"
- ii. "Can you help find the Froodles? They need to board the spaceship before it flies away. Drag the Froodles to the spaceship!"
- iii. "Some more Froodles are in here! Can you drag all the Froodles to the spaceship?"
- iv. "Find the Froodles! Put them on the spaceship as fast as you can!"

2. Yook

a. Prompts:

- i. "This is a Yook. Yooks have blue spikes and are round. Put this Yook on the spaceship!"
- ii. "There are more Yooks in here. Move all the Yooks to the spaceship!"
- iii. "There are still Yooks in this room! Put all of the Yooks on the spaceship!"
- iv. "Here's another room with aliens. Please put the Yooks on this spaceship!"

3. Micket

a. Prompts:

- i. "This is a Micket. You can tell because Mickets are purple, shaped like triangles, and have white spikes. Move this Micket to the spaceship!"
- ii. "Find all of the Mickets and put them on this spaceship!"
- iii. "There are more Mickets in here. Drag them all to the spaceship!"
- iv. "Still more Mickets in this room. Move all of the Mickets onto the spaceship!"

APPENDIX E

Study 2 Results Tables

Table 3.9. Feedback Error Analyses

	Incorrect Answers (Errors) by Feedback Group				
	Summary Statistics		ANCOVA Model without Interactions		
	<i>M</i>	<i>SD</i>	F-statistic	Contrast	p-value
Feedback Groups (N=240)			2.43		0.09
<i>Sounds (n=80)</i>	1.48	1.33			
<i>Encouragement (n=80)</i>	1.20	0.88			
<i>Scaffolding (n=80)</i>	1.11	0.92			
Covariates					
<i>Age</i>			6.11		0.01*
<i>Tablet Experience</i>			1.00		0.37
<i>Order of Play</i>			3.25		0.07
Comparisons (Tukey)					
Feedback Groups					
<i>Sounds v. Encouragement</i>				-0.27	0.21
<i>Sounds v. Scaffolding</i>				-0.33	0.10
<i>Encouragement v. Scaffolding</i>				-0.06	0.93
Tablet Experience Groups					
<i>No Experience v. Some Experience</i>				-0.17	0.76
<i>No Experience v. Very Experienced</i>				-0.29	0.40
<i>Some Experience v. Very Experienced</i>				-0.13	0.64

Condition: Feedback First

	<i>M</i>	<i>SD</i>	F-statistic	Contrast	p-value
Feedback Groups (N=120)			5.46		0.01**
<i>Sounds (n=40)</i>	1.71	1.22			
<i>Encouragement (n=40)</i>	1.30	0.69			
<i>Scaffolding (n=40)</i>	1.12	0.61			
Covariates					
<i>Age</i>			0.04		0.82
<i>Tablet Experience</i>			6.20		0.01**
Comparison (Tukey)					
Feedback Groups					
<i>Sounds v. Encouragement</i>				-0.45	0.06

<i>Sounds v. Scaffolding</i>	-0.60	0.01**
<i>Encouragement v. Scaffolding</i>	-0.16	0.70
Tablet Experience Groups		
<i>No Experience v. Some Experience</i>	-0.58	0.09
<i>No Experience v. Very Experienced</i>	-0.91	0.00*
<i>Some Experience v Very Experienced</i>	-0.33	0.13

Table 3.10. Feedback Response Times Analyses

	Response Times (RTs) by Feedback Group				
	Summary Statistics		ANCOVA Model without Interactions		
	<i>M</i>	<i>SD</i>	F-statistic	Contrast	p-value
Feedback Groups (N=240)			3.44		0.03*
Sounds (<i>n</i> =80)	11.13	4.86			
Encouragement (<i>n</i> =80)	13.65	8.64			
Scaffolding (<i>n</i> =80)	12.16	5.94			
Covariates					
<i>Age</i>			13.41		0.001***
<i>Tablet Experience</i>			2.27		0.11
<i>Order of Play</i>			19.30		0.000***
Comparisons (Tukey)					
Feedback Groups					
<i>Sounds v. Encouragement</i>				2.58	0.03*
<i>Sounds v. Scaffolding</i>				1.29	0.39
<i>Encouragement v. Scaffolding</i>				-1.29	0.39
Tablet Experience					
<i>No Experience v. Some Experience</i>				1.05	0.74
<i>No Experience v. Very Experienced</i>				-0.79	0.84
<i>Some Experience v. Very Experienced</i>				-1.84	0.09
Condition: Feedback First					
	<i>M</i>	<i>SD</i>	F-statistic	Contrast	p-value
Feedback Groups (N=120)			2.42		0.09
<i>Sounds (n=40)</i>	12.65	5.36			
<i>Encouragement (n=40)</i>	18.18	10.22			
<i>Scaffolding (n=40)</i>	13.48	5.82			
Comparison (Tukey)					
Feedback Groups					
<i>Sounds v. Encouragement</i>				3.41	0.08
<i>Sounds v. Scaffolding</i>				1.12	0.75
<i>Encouragement v. Scaffolding</i>				-2.28	0.32
Tablet Experience					
<i>No Experience v. Some Experience</i>				1.58	0.77
<i>No Experience v. Very Experienced</i>				-2.35	0.55
<i>Some Experience v. Very Experienced</i>				-3.93	0.01*

Table 3.11. Leveling Error Analyses

	Incorrect Answers (Errors) by Leveling Group				
	Summary Statistics		ANCOVA Model without Interactions		
	<i>M</i>	<i>SD</i>	F-statistic	Contrast	p-value
Leveling Groups (N=240)			10.52		0.01**
Scaffolded (<i>n</i> =120)	2.54	1.96			
Random Order (<i>n</i> =120)	3.34	1.96			
Covariates					
<i>Age</i>			1.98		0.16
<i>Tablet Experience</i>			0.01		0.99
<i>Order of Play</i>			12.20		0.001***
Comparison (Tukey)					
Tablet Experience					
<i>No Experience v. Some Experience</i>				-0.04	0.99
<i>No Experience v. Very Experienced</i>				-0.00	1.00
<i>Some Experience v. Very Experienced</i>				0.03	0.99

Table 3.12. Leveling Response Times Analyses

	Response Times (RTs) by Leveling Group				
	Summary Statistics		ANCOVA Model without Interactions		
	<i>M</i>	<i>SD</i>	F-statistic	Contrast	p-value
Leveling Groups (N=240)			13.92		0.00***
Scaffolded (<i>n</i> =120)	17.17	8.62			
Random Order (<i>n</i> =120)	21.51	10.88			
Covariates					
<i>Age</i>			12.52		0.00***
<i>Tablet Experience</i>			0.54		0.99
<i>Order of Play</i>			11.02		0.00**
Comparison (Tukey)					
Tablet Experience					
<i>No Experience v. Some Experience</i>				2.10	0.60
<i>No Experience v. Very Experienced</i>				1.21	0.83
<i>Some Experience v. Very Experienced</i>				-0.89	0.78