The Pennsylvania State University

The Graduate School

College of the Liberal Arts

# INFLUENCES ON CHILDREN'S PLAY WITH A STEM TOY: INTERACTIONS AMONG CHILDREN, PARENTS, AND GENDER-BASED MARKETING

A Dissertation in

Psychology

by

Emily F. Coyle



Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2015

The dissertation of Emily F. Coyle was reviewed and approved* by the following:

Lynn S. Liben
Distinguished Professor of Psychology, Human Development & Family Studies, and Education
Dissertation Advisor
Chair of Committee

Janet K. Swim
Professor of Psychology

Eric Loken
Research Associate Professor of Human Development & Family Studies

Carla Zembal-Saul
Professor of Science Education

Margaret L. Signorella
Professor of Psychology and Women's Studies
Penn State, Greater Allegheny

Melvin M. Mark
Professor of Psychology
Head of the Department of Psychology

*Signatures are on file in the Graduate School

**ABSTRACT**

Women and girls are underrepresented in science, technology, engineering, and mathematics (STEM) at most levels. Gender Schema Theory suggests that early on, girls and boys gravitate toward different experiences, approaching what they perceive as culturally appropriate for their gender and avoiding what they perceive as inappropriate. Gender-differentiated play may thus shape actual skill development and gender attitudes. Marketing toys as "for girls" or "for boys" may influence children's willingness to play. Because spatial/mechanical toys are traditionally marketed primarily to boys, and spatial skill has been linked to success in STEM fields, one contributor to the STEM disparity may be boys' early spatial skill advantage.

Although girls play less with construction toys, we know little about how children's and parents' gender beliefs influence play, and whether marketing affects toy interest. I examined these factors, focusing on mothers' and children's interest in, and play with, a construction toy marketed explicitly to girls versus to boys. *GoldieBlox* is a new toy with an accompanying story. Children address challenges that arise in the narrative by assembling simple machines (e.g., belt-drives), combining a traditionally feminine mode of play – reading – with a traditionally masculine one – construction.

Sixty-one mother-child dyads (age 4-6 years) played together with either *GoldieBlox* or *BobbyBlox* (modified to use masculine colors and male character). Mothers were given a few minutes to prepare for play and familiarize themselves with the toy. Dyads then played together for 15 minutes after which children played independently with the toy for 5-7 minutes. Children were assessed for mechanical learning at post-test.

Findings confirmed that toy marketing affected how mothers and children used the toys, what they thought about the toys, and how much children learned from the toys. Mothers engaged in more building during familiarization with *BobbyBlox* than with *GoldieBlox*. Mothers used different scaffolding strategies with daughters versus with sons, reading to daughters and demonstrating building to sons, resulting in differential support of mechanical learning. Girls built more overall with *GoldieBlox* and boys built more overall with *BobbyBlox*. Ultimately, however, girls learned more from *BobbyBlox* than from *GoldieBlox* and boys learned more from *GoldieBlox* than from *BobbyBlox*. These effects also interacted with individual differences in children's gender salience.

This research tested the role of the gender salience filter in Liben & Bigler's (2002) Dual Pathways model. Highly salient girls especially benefitted from play with *BobbyBlox* and did not learn when playing with *GoldieBlox*. This suggests that for STEM intervention, marketing STEM by gender may not be a successful approach. Overall, marketing powerfully affected toy use and learning, a finding that has both practical and theoretical implications.

# TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship, awarded early in my graduate training under Grant No. DGE1255832. Any opinions, findings, and conclusions or recommendations expressed in this material are my own and do not necessarily reflect the views of the National Science Foundation. This work was also funded in part by the College of the Liberal Arts RGSO dissertation award and a dissertation award made possible through the Child Study Center's Robert and Ruth Faris Child Psychology Fund in the College of the Liberal Arts. The Child Study Center has been generous in their support of my research, my training, and my conference travel and I am very thankful.

I am perhaps forever indebted to my advisor, Lynn Liben, for five years of tireless mentorship. I came to graduate school with questions and enthusiasm, and she has shepherded me down the path to forging an innovative, independent research agenda and career trajectory. She has been a constant support during some extremely difficult periods in the last several years, for which I am incredibly thankful. Although I am somewhat nervous to embark on my own, I am also excited because I know Lynn's mentorship is for life and in that regard, I am not really leaving at all.

I am deeply grateful to my undergraduate advisor, Megan Fulcher, who continues to be one of my most important mentors and collaborators. As my career has evolved, so too has our relationship – from advisor, to collaborator, to mentor and friend. I have most recently discovered that she is my working model for integrating research and teaching, and I draw on her insights constantly in the classroom. I look forward to many more years of mentorship and collaboration.

I would like to thank my doctoral committee, Janet Swim, Eric Loken, Carla Zembal-Saul, and Peggy Signorella, for their helpful feedback at every stage of this process. Together they are a truly interdisciplinary group with a multitude of strengths, and my research is (and I am personally) all the better for it.

None of this work would be possible without the parents and children who participated, or the parents, children, and preschools who assisted in the pilot phases of this work. Thank you in particular to Gail Guss at the Bennett Family Center and Linda Reichert at the Child Care Center at Hort Woods, both at Penn State, Laurel Sanders at Grace Lutheran Preschool, as well as Matt Neely, Jenny Cummins, and Alan Barta at Epiphany School. Truly, child development research is for the benefit of the child, and would not exist without families' enthusiastic participation.

I would be remiss not to thank the administrative and support staff that keep things running in Moore Building. In particular, Sherrilee Gilliland helped me to navigate the deadlines, communication, and paperwork that accompany the many graduate school benchmarks. Marc Kepler provided tireless technical assistance and creative solutions to solve my software and hardware challenges.

I am grateful to Cognitive and Social Development labmates, past and present, including Lacey Hilliard, Adam Christensen, Cori Bower, Giulia Borriello, and Kingsley Schroeder. Thank you for your thoughtful assistance at all stages of the process but most especially, for being my friends and supporters.

I would also like to thank the research assistants of the Cognitive and Social Development lab for their tireless assistance in designing study materials, recruiting participants, entering data, transcribing videos, coding videos, and generally being good sports about the entire research process. I am particularly indebted to Kevin Peno and Kitty Lam and thank them for their dedication, without which I might not have a dissertation at all.

I am incredibly grateful to my graduate school colleagues and friends. We have shared triumphs and at times, tribulations. I am thrilled to see what each of you will go out in the world to do. I am especially thankful for my dear friend Paige Deckert. She has been the source of thoughtful professional feedback, occasional personal criticism, and a veritable boatload of support.

I would like to thank my parents, Sherilyn Peterson and Richard Coyle, for their constant support, encouragement, and faith throughout my now 24 years of formal education. My mother in particular has been an inspirational example of just how hard one can possibly work, and just how high one can possibly reach. My baby brother, Maxwell Coyle, has been my number one cheerleader and I hope to return the favor when he decides on his own brilliant future. I am also grateful to my sister, Rachael Coyle, for enabling the cooking and baking habits that have kept me sane through the last five years.

I am grateful to my niece and nephew, Amelia Francis and Wynn Logan, for inspiring my work on a daily basis. I cannot wait to live fewer than 3,000 miles away so they can inspire me in person more often than on the phone.

Last but not least, I am grateful to my cat, Kiki Marie, and my horses past and present, Best Coast, Beg to Differ, Cheap Thrills, and Hershey Kiss. Though they cannot actually read this and Kiki might eat the paper, they all deserve credit for their constant and nearly judgment-free support. Without their presence to keep a rhythm in my life, graduate school surely would have driven me mad. I love you all.

# Chapter 1

## INTRODUCTION

Women and girls are underrepresented in science, technology, engineering, and math domains (STEM), a gender gap that begins early and persists throughout the lifespan. Although girls receive higher grades than boys in math and science from elementary school to college and although a recent meta-analysis found no difference between males and females in math performance (Lindberg, Hyde, Petersen, & Linn, 2010), women are underrepresented in mathematically intensive college majors and in STEM fields after college (see Ceci & Williams, 2010 for a review). Given that spatial skill has been linked to success in STEM fields (e.g., Shea, Lubinski, & Benbow, 2001; Sorby, Casey, Veurink, & Dulaney, 2013; Wai, Lubinski, & Benbow, 2009; Wolfgang, Stannard, & Jones, 2003), one contributor to this disparity may be an early and persistent gap between boys and girls in spatial skill, with boys outperforming girls (e.g., Halpern et al., 2008; Linn & Petersen, 1985; Uttal et al., 2013).

This gap in spatial skill and the tendency to hold gender stereotyped career preferences begin very early, even as young as preschool (e.g., Fulcher, Sutfin, & Patterson, 2008; Helwig, 1998; Liben & Bigler, 2002). Early play, before formal education even begins, may be a critical opportunity for building STEM-relevant skills and interests (Wolfgang et al., 2003). Boys and girls generally play with different toys and in single-sex peer groups (e.g., Cherney, Kelly-Vance, Gill Glover, Ruane, & Oliver Ryalls, 2003; Maccoby & Jacklin, 1974; Martin, Fabes, Hanish, & Hollenstein, 2005; Martin et al., 2012), a pattern that might result in differential exposure to opportunities to build STEM-relevant skills in early play.

Many formal interventions to combat the gender gap in STEM have been targeted to a variety of age or education levels and in a variety of contexts. Such programs are delivered in classrooms, as part of extra-curricular groups like the Girl Scouts, in university contexts, or in laboratory settings, among others (see Liben & Coyle, 2014 for a recent review). Popular toys have not been empirically evaluated as STEM interventions, although play in the form of storytelling has been used successfully to teach math skills in early grade school (e.g., Casey, Kersh, & Young, 2004). Indeed, much remains to be understood about the earliest precursors to girls' long-term lower interest in STEM fields relative to boys. For example, although girls play less with spatial and construction toys than do boys (e.g., Cherney et al., 2003; Jirout & Newcombe, 2015), the extent to which this gender difference is child-driven, parent-driven, peer-driven, or an artifact of gender-based marketing remains unknown. It is important to disentangle these influences to understand the antecedents to girls' reduced interest in STEM and to identify what will likely be successful in intervention to increase STEM participation. As well, testing these influences empirically will contribute to constructivist theoretical development including gender schema theory and the role of the individual differences as moderators of gender development (Liben & Bigler, 2002).

The current study begins to address these issues by focusing on three related questions: (1) Do girls and boys approach and use STEM toys differently, irrespective of whether those toys are explicitly marketed or designed to appeal to girls versus boys? (2) Do parents guide play with a STEM toy differently with daughters versus with sons? (3) How is play affected if the toy is marketed explicitly to girls versus to boys? The current study examines these questions in the context of a new engineering toy marketed explicitly to girls, *GoldieBlox*, discussed in more detail later in this proposal.

**Gender Development and STEM**

According to a schema theory perspective on gender development, boys and girls seek out information about their own group (Martin & Halverson, 1981). They approach what they perceive as culturally appropriate for their own gender and avoid what they perceive to be inappropriate. This information is consolidated into “gender schemas,” a framework for understanding gender, guiding personal gender behavior, and predicting others’ gender behavior. Early gender schema theorists hypothesized that individuals differ in their relative tendency to use gender as an organizing principle in cognition and behavior (e.g., Bem, 1983). Liben and Bigler (2002) expanded schema theory to give more explicit attention to the role of individual differences in specific interests and personal gender salience. According to their dual pathways model, both of these factors are important moderators of children’s developing gender schemas because they influence actual approach or avoidance behavior. For example, for a child to approach a situation, he or she must be personally interested in the situation. In order for gender schemas to be influenced by that engagement, the child must perceive gender to be a salient dimension on which to categorize. That is, gender must be personally important or cognitively relevant. If gender is personally salient and if the child endorses traditional gender values, the child must also perceive the situation to be gender-appropriate in order to approach. Otherwise, the child will ignore the situation. For example, a girl who uses gender as an important guiding principle in her approach or avoidance behavior will approach playing with a construction toy if she is both interested in the toy and if she perceives the toy to be appropriate for girls.

**Personal gender salience.** Personal gender salience, termed the *gender salience filter* (GSF) in Liben and Bigler’s model, is an especially important individual differences variable. As touched on already, personal gender salience refers to the degree to which gender is personally relevant in decision-making about behavior and in guiding cognition. The GSF has been measured in previous work (e.g., Coyle & Liben, 2013) using multiple dimensions that tap the degree to which children's cognitive focus is steered or directed by gender: (1) self-perceived gender typicality, (2) gendered affiliative preference for peers, and (3) incidental memory for gender. Self-perceived gender typicality, the degree to which a child feels he or she is like same-sex versus other-sex peers, assesses the extent to which a child aligns behavior and belonging on gender lines. Gendered affiliative preference for peers assesses whether behavior preferences are guided by gender. Incidental memory for gender assesses whether attention generally is directed to gender. These three measures capture two related concepts in gender salience: cognitions about typicality and behavior, and attention to gender more broadly. Experimental work finds that highly gender salient children are especially affected by the presence of gender-related cues (Coyle & Liben, 2013).

More specifically, Coyle and Liben (2013) had preschool girls participate individually in a computer game about masculine, feminine, and novel occupations (Liben, Bigler, & Krogh, 2001), with either a strongly feminized character or a weakly feminized character. The strongly feminized character had a more stereotypically feminized physical appearance (specifically, a Barbie doll with a narrow waist and large breasts) than did the weakly feminized character (a female Playmobil doll with an average waist and no defined breasts). Highly gender-salient (H-GSF) girls showed an increase in feminine interests after playing the game with the strongly feminized character. This suggests that the feminized gender cues in the game reinforced H-GSF girls' existing feminine interests. There was, however, no significant reinforcement of feminine interests among these girls when playing the game with the weakly feminized character. In contrast, low gender-salient (L-GSF) girls showed no change in their level of traditionally feminine interests after playing with either the strongly or weakly feminized character. Regardless of condition and GSF strength, girls' interest in traditionally masculine and feminine jobs remained unchanged. These results highlight the important and interactive role that personal gender salience has with gender-relevant encounters on children's developing gender schemas (Coyle & Liben, 2013).

It is clear from previous work that personal gender salience is an important, although often overlooked, individual-difference variable affecting children's cognitions, behavior, and potentially their gender schema development. Gender schemas are reinforced or modified after a gender-relevant encounter for H-GSF children. In fact, H-GSF children may perceive gender to be relevant in myriad encounters by virtue of their tendency to categorize using gender. As this applies to developing STEM interests and abilities, H-GSF girls especially may not approach situations that might hold promise for developing interest or ability because they perceive the situation as gender-relevant but gender inappropriate. Research suggests that even the ambient environment in a STEM classroom feels gender inappropriate for some girls (e.g., Cheryan, Meltzoff, & Kim, 2011; Cheryan, Plaut, Davies, & Steele, 2009). When the environment is manipulated to be more gender neutral, girls are more interested and successful in STEM tasks. But even contexts designed to be obviously gender appropriate (i.e., highly feminized) may backfire as interventions for increasing girls' participation in STEM. This may be the case if the gender context of the situation reinforces existing gender stereotypes instead of inducing schema modification (e.g., Murphy & Whitelegg, 2006; Weisgram & Bigler, 2006, 2007). Therefore a girl's approach to STEM contexts may be in part driven by the encounter itself and in part by the individual, according both to interest and gender schematic beliefs.

**Gender in toy marketing.** Another situation with strong cues about gender appropriate behavior is play. Toys marketed to girls versus to boys are different in color and content, and girls and boys approach them in different numbers (e.g., Auster & Mansbach, 2012; Blakemore & Centers, 2005; Cherney et al., 2003; LoBue & DeLoache, 2011). Given that most spatial and construction toys are marketed to boys (Auster & Mansbach, 2012), very young girls might not approach and play with construction toys, consequently failing to develop the associated skills or interest that could lead to future STEM encounters. In contrast, a construction toy marketed to girls might successfully induce approach and play, especially among H-GSF girls. The importance of the dual role of child-driven approach and toy-driven appeal become especially clear in this example. Why do girls not just play with construction toys already on the market, those that are geared primarily toward boys? Likely there are many reasons for this, including parent-driven and peer-driven influences on play behavior. Certainly there is child-driven avoidance of a perceived gender inappropriate toy, according to the gender schema theory principles already discussed. With toys, especially, children may show a “hot potato” effect leading them to avoid gender inappropriate toys. Martin, Eisenbud, and Rose (1995) gave preschool children novel toys and children rate how much they personally liked it, how much boys would like it, and how much girls would like it. Children predicted their same-sex peers would like toys they rated highly but that other-sex peers would not like those toys. When novel toys were labeled as explicitly “for girls” or “for boys,” children showed an especially large difference in their ratings, being most interested in toys with their same-gender label and least interested in toys with the other-gender label. Indeed, boys and girls showed less interest in other-gender labeled toys regardless of the attractiveness of the toy, suggesting they are not interested even in highly attractive toys because of the gender label (the hot potato effect). As well, girls were especially influenced by explicit labeling, giving even higher ratings to same-gender labeled toys than did boys (Martin et al., 1995).

Explicit labels are one means of indicating gender appropriateness, and color is another. LoBue and DeLoache (2011) examined the developmental preference for or avoidance of pink in particular. They gave children groups of novel objects, with one object in each group being pink and the others being the same but not pink. By 36-months-old, girls selected the pink object more often than chance, and boys selected the pink object much less often than chance. This difference became especially pronounced by the preschool years, with girls showing a dramatic increase in their selection of pink by three-years-old, and boys showing a greater avoidance of pink by four-years-old (LoBue & DeLoache, 2011). It is well documented that by preschool, boys and girls show a tendency to play with different, sex-typed toys (Cherney et al., 2003).

Newer research integrates explicit labeling with implicit color cues. Fulcher, Weisgram, and Dinella (2013) gave preschoolers novel toys that were either pink or blue, and told children they were either "for girls" or "for boys." Children rated their own and others' interest in the toys. For boys, the explicit label was most important; boys were interested in either pink or blue toys labeled "for boys" and less interested in pink or blue toys labeled as "for girls." Girls showed the hot potato effect strongly, being influenced both by explicit labels and toy color. Girls were most interested in anything labeled "for girls," regardless of color, rating novel toys even higher than boys did. But, girls were less interested in pink toys labeled as "for boys" and dramatically less interested in blue toys labeled "for boys" (Fulcher et al., 2013). Taking all of these studies together, it is clear that gender cues influence young children's interest in toys, although girls may be especially attuned both to explicit cues like labeling and implicit cues like color.

Apart from whether toys are marketed primarily to girls or to boys, toys carry with them potentially influential messages about gender. They may implicitly teach or model what is gender-appropriate by virtue of being marketed as “for girls” or “for boys.” Therefore it is important to understand how toys and media can be influential, impacting children’s developing gender attitudes. The developmental media persuasion model predicts a three-level model of the persuasive nature of media, including toys, on children’s attitudes and beliefs (Buijzen, Van Reijmersdal, & Owen, 2010). Although typically applied to explain the influence of media on children’s attitudes towards a particular commercial product, this model is also helpful to understand how media and toys are influential in shaping developing gender attitudes. Systematic-persuasion processing occurs when a child deliberately thinks about and elaborates on a message. This is the deepest level of processing and also the least likely level to occur among young children. Heuristic-persuasion processing involves moderate attention to persuasive messages, guided primarily by heuristic cues like the attractiveness of the toy or packaging. The most surface-level processing, automatic-persuasion processing, occurs with little attention to a particular message and instead, the response is driven by other cues, for example how much fun a child has while playing with a toy. Young children are most likely to process media information at this level (Buijzen et al., 2010). Toys that are fun may have more impact on children’s attitudes and behaviors. Children are likely to play repeatedly with these toys and to internalize the associated messages about gender, for example believing that toy, and similar toys, to be gender-appropriate. Attractive toys in colorful packaging may be especially appealing. The influence of toy attractiveness and fun suggest that toys themselves may attract a child to approach and play. Additionally, children drive approach or avoidance based on perceptions of gender appropriateness, like color or explicit label.

**Parents, Gender Development, and STEM**

Parents are another important source of information about gender. Research shows that parents' own gender attitudes as well as their behavior in and out of the home are predictive of children's gender beliefs (e.g., Fulcher, 2011; Tenenbaum & Leaper, 2002). Because young children spend a great deal of time with parents, parents may be particularly influential early in development. Indeed, cross-sectional research suggests that parents' attitudes and behavior differentially influence children's developing gender cognitions across childhood, adolescence, and emerging adulthood (Fulcher & Coyle, 2011). Early in life, visible behavior, such as division of household labor and participation in paid labor, may be especially influential. Across development, children also assimilate parents' attitudes into their own gender schemas. Emerging adults' own gender beliefs moderate the connection between parents' attitudes and behaviors, and adult children's own gender attitudes and plans for work and family (Fulcher & Coyle, 2011).

Fulcher et al. (2008) studied a sample of heterosexual and lesbian families to examine the relationship between (a) parents' attitudes about gender and their division of paid and unpaid labor and (b) preschool children's knowledge of stereotypes and their sex-typed activity preferences. More conservative parents (as measured on gender attitude assessments) had children who believed gender transgressions were more serious and they themselves had more gender stereotype knowledge and more highly gender-stereotyped occupational aspirations. Parents with a more unequal division of paid labor had children with greater sex-typed preferences for activities; parents with more unequal division of household labor had children with more traditional career aspirations and greater gender stereotype knowledge. Division of childcare labor was the best predictor of children's career aspirations, where more egalitarian families had children with the least sex-typed aspirations. Fulcher (2011) also found a similar association between parent behaviors and attitudes and child self-efficacy in a variety of domains. Mothers with less gender-traditional attitudes, less gender-traditional occupations, and more egalitarian division of household labor with partners had children with greater self-efficacy in nontraditional domains, including STEM.

Some of parents' influence on children's beliefs and behavior may derive from parents' active socialization of gender in their children (e.g., Katz, 1996, Blakemore & Hill, 2008). Active socialization can mean directing behavior towards gender appropriate activities or punishing gender inappropriate behaviors perceived to be transgressions. The extent to which this occurs is in accordance with parents' own gender attitudes; less traditional parents might socialize behavior very differently. Katz (1996) offered an informal model for how feminist parenting specifically might be carried out, in light of the association between division of labor and child gender flexibility. She argued that different sources of gendered information are important at different developmental stages. Before six, parents and toys may be most important, while teachers, peers, and access to media become relevant during elementary school (Katz, 1996). Blakemore and Hill (2008) developed a scale for testing the extent to which parents are trying to socialize their children in the domains that Katz discusses. Thus parent influences on child attitudes about gender and STEM may be both parent-driven, as with active socialization or exposure to particular STEM-relevant experiences, and child-driven, as with assimilating parent attitudes and behaviors into a child's developing gender schema.

Simpkins, Fredricks, and Eccles (2012) explored how these parent-driven influences operate ultimately to become child-driven influences in a developmental cascade. They examined mothers' beliefs about children's ability and efficacy and mothers' skill-promoting behavior in four areas: reading, math, sports, and music. They looked at children's beliefs about their own ability, valuing of activities, participation, and actual ability in a cross-sequential design that tracked development from as young as kindergarten to potentially as old as twelfth grade. Unsurprisingly, mother's beliefs about a domain influenced her behavior. More skill-promoting behavior in mothers resulted in more positive beliefs about that domain in children. Children's positive beliefs were associated with greater participation in that domain. Indeed, mothers' beliefs and behavior mediated children's valuing of a domain, evidence that parents' beliefs and actual behaviors jointly influence children's cognition and behaviors (Simpkins et al., 2012). This model supports the Expectancy-Value framework that predicts interest and success in a domain are a product of expectations about success and valuing of the domain (e.g., Eccles, 2014; Eccles et al., 1983)

Importantly, Simpkins and colleagues found that mothers provided different skill-building, behavioral support to girls versus boys (Simpkins et al., 2012). Mothers provided more opportunities and encouragement for boys in sports and for girls in music. As well, mothers believed boys and girls were differently able in these domains. This is an important demonstration of how parents' beliefs about gender shape the skill-building opportunities they provide for sons and daughters. Extending this argument in the context of the expectancy-value framework to consider early play and STEM, it is possible that if parents value STEM differently for sons versus daughters or have different expectations about their abilities, they might provide very different opportunities for skill development. Child learning might therefore vary accordingly by gender. For example, parents who hold stereotyped beliefs might purchase construction toys for sons but not for daughters. The parent-driven influence is therefore critical to consider when answering questions about early STEM interest and approach. They are not, however, the only influence on play behavior outside of the child. Peers, potential playmates or classmates, are also an important proximal influence on gendered behavior and cognition.

**Peers, Gender Development, and STEM**

Peers are a well-established influence on child social, emotional, and cognitive development (e.g., Rubin, Bukowski, & Parker, 2006). They may be influential at the level of interactions between two children, relationships or bonds between children, and at the level of the group or collection of peers. Peers can influence a child's actual behavior and their cognitions. Some have suggested peers may be even more influential on development than are parents (Harris, 1995, 2009). Meta-analysis (Degner & Dalege, 2013) suggests the more tempered conclusion, that parents and peers both matter, but certainly we cannot discount the role of peers in development, including in gender development. A child's peer group, from early in life, is generally same-sex (Maccoby & Jacklin, 1974; Martin et al., 2005; Martin et al., 2012). Martin and colleagues (2005) observed preschoolers' interactions at play and in the classroom over one year. They used the state-space grid approach based on dynamic systems principles to model interactions over time. Same-sex peer interaction states were more common than mixed-sex states, meaning most interaction occurred with same-sex groups. Additionally, children who engaged in more same-sex play showed an increase in preference for play with same-sex peers, so that same-sex play for these children became increasingly common. The authors conclude that there are mechanisms working across time to promote same-sex play (Martin et al., 2005). Martin and colleagues (2012) used a similar methodology and stochastic actor-based modeling (SAB) to examine the influence of the group outside of the target interaction partners. Again they found children played with same-sex peers more often, and with peers with similar activity preferences. Importantly, they found the play preferences of the group were also powerful in shaping a child's activities and interactions (Martin et al., 2012). In addition, groups of children may police one another's gendered behavior and punish those behaviors perceived to be gender transgressive (e.g., Blakemore, 2003; Smetana, 1986). A climate of same-sex groups with gender traditional modes of play may be self-perpetuating.

Peers are influential at the level of actual activities (e.g., engaging in gender traditional modes of play within same-sex play groups), as well as at the level of cognitions (e.g., what children actually think about gender and group membership). Martin and colleagues (2011) examined the association between children's actual play behavior and their gender attitudes and beliefs about self-perceived similarity to same versus other-sex peers. They assessed self-perceived similarity implicitly by asking children to rate their own interest in a novel toy as well as how much another boy or a girl would like it. Children who were more similar to same-sex peers made similar interest ratings for themselves and their same-sex ingroup, but made a different rating for the other-sex outgroup. In fact, children's playmate preferences were better predicted by perceived similarity to peers than by congruence with peer activity preferences. And, perceived similarity increased with age, suggesting that playing with same-sex peers reinforced cognitions about self-perceived similarity (Martin et al., 2011). Thus H-GSF children who perceive themselves as more gender typical are most likely to play with same-sex peers and in gender traditional modes of play. As this relates to STEM and early play, H-GSF girls may be more likely to engage in traditionally feminine play with same-sex peers, building those related skills and interests. They are more likely to miss out on opportunities like construction play and consequently, to miss out on associated antecedent skills needed to succeed in STEM.

Early play is one means of building concrete skills and interests, as discussed already, and peers are likely strongly influential in early play experiences. This may be peer-driven, by the larger group climate or the individual peer's direction, or child-driven, by their engagement in same-sex play groups, gender traditional play, and the assimilation of all of that into developing gender schemas. Skills or interests built in play are honed in later formal education, an arena in which peers are also influential. The peer group influences academic self-concept and engagement in formal schooling (Molloy, Gest, & Rulison, 2011). Peers who are more engaged in school and in STEM domains may be supportive for children to stay engaged in STEM, whereas peers who are disengaged from STEM domains may prompt disengagement in the child as well.

**Existing Empirical Tests of STEM Intervention**

So far I have discussed particular child-driven, parent-driven, and peer-driven influences on gender development and approach and avoidance behavior predicted by gender schema theory. This applies well in the domain of girls' participation in STEM because gender is deeply intertwined in perceptions of STEM. I will now discuss examples of existing intervention approaches to increasing girls' participation in STEM. Because increasing girls' interest and participation in STEM domains is currently considered to be a critical social, economic, and political goal, intervention has been approached in many ways. Generally, interventions to increase girls' participation or interest in STEM operate on one of the two principles in Liben and Bigler's (2002) dual pathways model: personal interest or gender cognitions. Few explicitly test this theory, but most can be categorized as intending to change girls' and women's own attitudes and stereotypes about STEM, change or expand their interests to include STEM, or to change girls' and women's cognitions about the links between their own interests and their beliefs about STEM. Intervention approaches vary in the extent to which they target girls or women (by attempting to expand or break stereotypes, or by expanding interest) versus target science (by attempting to showcase how science is not stereotypically masculine or can be stereotypically feminine). Liben (2012) proposed a more formal taxonomy for categorizing how intervention has been approached so far, and Liben and Coyle (2014) refined and expanded this taxonomy in their recent review. For organizational purposes, I will use their taxonomy to classify exemplars of existing STEM interventions. The five categories are: Remediate, Revise, Refocus, Recategorize, and Resist.

*Remediate* approaches to intervention for girls and STEM involve modifying girls by building relevant skill sets and related self-constructs. Such approaches might focus on developing spatial skills or building self-efficacy for science. Indeed, spatial skills have been found to be malleable (see Uttal et al., 2013 for a meta-analytic summary) in some contexts, so this may be an effective point for intervention. De Lisi and Wolford (2002) used familiar children's computer games to teach spatial skills to third graders. Tetris was intended to provide experience and practice with mental rotation, "Where in the USA is Carmen Sandiego?" was intended to provide practice with social studies without providing mental rotation practice. Boys and girls played the games as part of a computer lab course. After one month, girls and boys who played Tetris improved in their mental rotation skills as compared to children who played Carmen Sandiego, and the gender gap in mental rotation vanished for the Tetris players (De Lisi & Wolford, 2002).

Of course, there are other STEM-relevant skills besides spatial skills. Mammes (2004) looked at German third grade boys and girls taught two units of technology education as compared to equivalent classrooms that did not receive this education. Children in the intervention classrooms first learned how to build electric circuits in parallel to make Christmas tree lights for a model Christmas tree. In the second unit, they learned about product development and production, from prototype development of a “nesting box” to the difference between batch and serial production. Before the intervention, girls in both the experimental and control classrooms had less interest in, less experience with, and less knowledge about technology than did boys. After six weeks, girls in the experimental classrooms increased in their interest, experience, and knowledge compared to girls in the control classrooms, and reduced although did not eliminate the gender gap (Mammes, 2004). Both of these studies are examples of teaching concrete, STEM-relevant skills as a means of increasing girls’ participation in STEM. Only Mammes (2004), however, examined actual change in interest.

*Revise* approaches to intervention aim to modify STEM to better match it to perceived feminine interests, preferences, or qualities. One example of such an approach comes from computer science, where interventionists have joined storytelling with programming to better attract girls (e.g., Carbonaro, Szafron, Cutumisu, & Schaeffer, 2010; Kelleher & Pausch, 2007; Kelleher, Pausch, & Kiesler, 2007; Lynn, Raphael, Olefsky, & Bachen, 2003). Kelleher and colleagues (2007) developed Storytelling Alice as a modification to the programming environment, Alice. Alice permits users to work entirely within a graphic interface, outside of the coding environment, making it approachable for novice coders. Whereas Alice allows users to program a humanoid character to move around, Storytelling Alice allows users to program a humanoid character to move, speak, and think, over several scenes. Middle school girls learned both Storytelling Alice and Generic Alice in a two-hour workshop. Although girls learned programming equally from both games, they were more interested in using Storytelling Alice again, were more likely to sneak extra time with Storytelling Alice, and spent more time actually programming with Storytelling Alice. The more traditionally feminine activity of building a story may be more successful in keeping girls engaged with computer programming over the long term even though both conditions were successful in teaching girls programming in the short term (Kelleher & Pausch, 2007; Kelleher et al., 2007). In fact, boys benefit from Storytelling Alice as well (Rodger et al., 2009).

Another *Revise* approach is to look at the ways that girls are made to feel they do not belong in science, and to address issues of exclusion by changing the STEM environment. For example, girls may not feel they belong in science because it is stereotyped as “nerdy.” Changing the ambient environment in which science is taught to be more neutral may feel more accepting. Cheryan and colleagues (2009) manipulated the objects in a classroom to be either stereotypical of computer science or more neutral. For example, the stereotypical classroom had a Star Trek poster, comic books, video games, and technical books. In the neutral classroom there was a nature poster, art, general interest books, and magazines. In one study, participants simply entered the classroom, sat for one minute, and then filled out surveys about computer science. Women, but not men, were less interested in computer science in the stereotypical classroom. In the neutral classroom, women were as interested as men in computer science (Cheryan et al., 2009). This small manipulation had important consequences for interest.

*Recategorize* approaches to intervention aim to replace beliefs about STEM and female gender as mutually exclusive with the possibility of multiple classification. At the most basic, a *Recategorize* intervention might be one that uses female scientist role models. Expanding Your Horizons (EYH) is a one-day program for middle school girls, typically held in university settings. Girls learn about STEM careers and are exposed to female scientist role models (Expanding Your Horizons Network, 2013). *Refocus* approaches to interventions emphasize correspondences between STEM and traditional femininity. For example, such approaches might emphasize the altruistic nature of conducting science, in order to show the existing correspondences between science and being traditionally feminine. Weisgram and Bigler (2006) attempted just this as a manipulation of the traditional one-day EYH program. Middle school girls were either assigned to the normal one-day EYH program (*Recategorize* intervention type) with female scientists describing their careers or to an altruism condition (*Refocus* intervention type), in which the scientists emphasized how their careers help people. Girls in this condition learned why science helps people and how scientists give back to others. They found no particular influence of the altruism emphasis on increasing girls' perception of science as altruistic or on their interest in science. But, belief in the altruistic value of science predicted interest in science, suggesting that recognizing altruism in science may indeed be important for girls, even if the EYH manipulation was not effective over and above the normal EYH program (Weisgram & Bigler, 2006).

One final set of *Recategorize* interventions of interest for the present study take multiple classification one step further by emphasizing co-occurrence of the unrelated identities of being female and being a scientist in a hyperfeminized scientist. Such approaches ignore or deny role conflict between girls' female and science identities while also idealizing hyperfemininity. The Science Cheerleaders are one example of this approach in the popular media (Cavalier, 2014). Science Cheerleaders are former National Football League cheerleaders now working in STEM fields. They put on cheerleading performances about science. On the one hand, the women embody the cultural ideal of feminized beauty in a traditionally feminine activity, cheerleading. On the other hand, they are actually employed as scientists. Their performances are displays of both of those simultaneous identities. Another example of feminized science models is Coyle and Liben's (2013) computer game work discussed earlier. Many of the jobs girls learned about in the computer game were STEM jobs, portrayed either by a strongly feminized character or a weakly feminized character. Although the strongly feminized character was most engaging for H-GSF girls, the game did not increase their interest in STEM careers pre- to post-test. Instead, the feminized model reinforced existing stereotypes about gender for H-GSF girls (Coyle & Liben, 2013). Importantly, this particular example would suggest such an approach might have unintended negative consequences in the domain of gender stereotyping. Other work suggests that affirming one's own feminine identity among potentially H-GSF girls or women is protective under conditions of stereotype threat, however (e.g., Miyake et al., 2010; Taillandier-Schmitt, Esnard, & Mokounkolo, 2012).

Clearly, intervention to increase girls' and women's interest and participation in STEM has been approached in myriad ways. The examples discussed are illustrative of the types of approaches taken to intervention but are certainly in no way exhaustive. However there is still extremely little empirical evaluation of the many intervention programs being undertaken, for example programs through schools, extracurricular groups, and camps.

**Play and STEM intervention.** Toys are another arena where science is being brought to girls on traditionally feminine terms (*Revise* intervention type) and play is potentially a critical point for intervention. Although it is well-established that boys and girls play in same-sex groups and with different toys from a young age, there is also a small but increasing movement to create toys marketed explicitly to girls that target STEM-relevant skills or interests. *LEGO Friends* are one such toy, similar to regular LEGOs except for their pink packaging, larger and more detailed figurines designed for make-believe play as well, and traditionally feminine themed sets. To my knowledge there is not yet any empirical evaluation of their effect.

*GoldieBlox* is another new toy explicitly marketed to girls in an attempt to increase interest and participation in STEM. *GoldieBlox* is a series of engineering construction sets with accompanying stories. Children address challenges that arise in the narrative by assembling simple machines, thereby combining a traditionally feminine mode of play – reading – with a traditionally masculine mode of play – construction. Because *GoldieBlox* is the focus of the present study, the design and mechanicals of the toy are described in more specific detail in the Method section. *GoldieBlox* has the potential to familiarize girls with construction play, something with which most girls are less involved. Additionally, it has the potential to increase actual mechanical skills given the engineering principles in the construction and play itself. Perhaps most importantly, it might result in girls becoming more interested in pursuing play with other toys that build STEM-relevant skills. LEGOs, for example, have been found to teach spatial skills. In one longitudinal study, children who showed more sophisticated play with LEGOs in preschool had stronger math achievement in middle school (Wolfgang et al., 2003). Puzzles, too, appear to promote spatial development. Levine, Ratliff, Huttenlocher, and Cannon (2012) observed everyday parent-child puzzle play in the home. Children who played more often with puzzles and with more difficult puzzles in preschool had better spatial skills than children who played with easier puzzles or played less often with puzzles. Although puzzles are not generally considered to be sex-typed, on average, boys played with more challenging, spatial skill-promoting puzzles than did girls (Levine et al., 2012).

## The Present Study

Although we know that girls play less with construction toys, we know little about how this play is affected by children's and parents' gender beliefs, and about whether toy interests and use can be affected by changing the way a toy is marketed. My dissertation examined these factors, focusing on parents' and children's interest in, and play with, a construction toy marketed explicitly to girls versus to boys. Behavior and interest were compared between children given *GoldieBlox* versus children given the same toy modified for boys (i.e., using masculine colors and a male character, "*BobbyBlox,*" discussed further in the Method). Toys in general may be effective in building STEM interest because play is fun. Enjoying oneself engenders positive attitudes towards a toy (Buijzen et al., 2010), and potentially toward an entire domain or subject. Playing repeatedly with a STEM toy could possibly build both interest in STEM and skills that are relevant to STEM, encouraging additional future STEM encounters. Given the gender gap in STEM and the lifelong gender gap in spatial ability, it is important to understand why boys and girls play with such different toys, potentially leading them ultimately to develop different skills and interests. As well, it is critical to understand how girls might be motivated to play with a toy that could potentially prompt development of STEM-relevant skills or interests given that skills built in early play appear to be important antecedents to later STEM-related skill (Levine et al., 2012) and achievement (Wolfgang et al., 2003).

The literature so far has established that boys and girls generally play in gender-segregated groups and with different toys. The logical assumption based on these observations and predicated on gender schema theory is that boys and girl consequently develop different interests and skills. It remains to be understood whether girls might be enticed to play with a traditionally masculine toy if it were presented as explicitly feminine. It is unknown whether boys and girls play with the same toys differently, for example by incorporating aspects of more masculine construction play versus more feminine make-believe play, and how parents guide play with construction toys with sons versus daughters. Therefore this study was designed to answer three closely-related research questions: (1) Do parents guide play with a STEM toy differently with daughters versus with sons? (2) Do girls and boys approach and use STEM toys differently, irrespective of whether those toys are explicitly marketed or designed to appeal to girls versus boys? (3) How is play affected if the toy is marketed explicitly to girls versus to boys? I am to begin to elucidate the earliest precursors to girls' reduced interest in STEM. To address the research questions, children and their mothers were observed playing with *GoldieBlox* or *BobbyBlox*.

The current study focuses on the role of parents because parents are children's gateway to toys and important primary playmates. Although I have also discussed the importance of peers in shaping children's gender development and play, peers are not included in the present study because practical limitations constrain the scope of this work. Parents are a critical first step in this work, and future research should examine the role of peers in children's interest in and play with STEM toys. Although both parents are likely important in shaping children's play, mothers were the participating parent in the laboratory session in present study for several reasons. First, mothers were much more likely to participate than were fathers. Second, mothers are typically responsible for the majority of childcare, even with historical shifts in division of household labor (Parker & Wang, 2013). Third, a consumer survey given to purchasers of *GoldieBlox* found that 73% were mothers purchasing for their children. Finally, there may well be gender differences between mothers and fathers that could influence their use of *GoldieBlox*, thus without the ability to test for those differences by having large and equal numbers of mothers and fathers, the sample was restricted to mothers. Although it was not possible to examine fathers' role proximally in the laboratory session, I actively solicited mothers' reports about fathers' role in the family to permit examining their role distally.

A secondary aim of this work was to test GST (Martin & Halverson, 1981), the dual pathways model (Liben & Bigler, 2002), and the moderating role of the GSF. Previous work has tested the role of the GSF in moderating girls' interest. The current study included boys as well as girls and examined both behavioral (i.e., play) and cognitive (i.e., interest and learning) outcomes. More generally, this research should speak to the larger issue of understanding what will be successful in gender and STEM intervention. Thus, this work has both theoretical and translational implications for a critical social problem, girls' underrepresentation in STEM, as is the goal of action research (Lewin, 1946).

**Pilot testing.** Informative pilot data came from (a) videos from *GoldieBlox Inc.* consumer testing, (b) YouTube videos parents posted online, (c) individual and small group testing I conducted with the toy, and (d) a consumer research survey conducted by *GoldieBlox Inc.*. Data from the first three sources demonstrated the range of variability in how children and mothers use the toy. Some children focused on reading the book or focused on make-believe play with the figures, whereas other children focused on construction. Videos of mothers playing with children showed a variety of strategies used to guide play. Some mothers particularly focused on reading the accompanying story, to the exclusion of building, a tactic that might foster reading skill but is unlikely to foster mechanical skill. Systematic empirical data are needed to better understand how parents guide play with a STEM toy and how this varies in relation to individual differences like gender attitudes or spatial skill.

The consumer research survey indicated that parents (73% were mothers) who purchased *GoldieBlox* during a pre-order period prior to the toy's initial release in January 2012 had a variety of motivations for doing so. Some indicated a desire to increase their daughter's STEM skill or interest, while others indicated a desire to teach their daughters about feminism generally. Motivation could be an important moderator of mothers' guidance in play, with mothers who are more STEM-focused giving more support to construction than parents who are less STEM-focused. These pilot data highlighted the importance of testing empirically how mothers use

STEM toys marketed to girls versus to boys, with daughters versus sons. The present study aimed to begin to disentangle the influences of parents, children, and toy marketing on play with a STEM toy.

**Hypotheses.** Although aspects of this work were exploratory in nature, a number of specific predictions were made with respect to children's and mothers' play with, interest in, and learning from *GoldieBlox* or *BobbieBlox*, as well as the role of individual differences in those outcomes. During familiarization, mothers were expected to engage in more construction with *BobbyBlox* than with GoldieBlox and more reading with GoldieBlox than with *BobbyBlox*. During dyadic play, mothers were expected to use more supportive scaffolding play strategies with sons than with daughters, although this was expected to vary depending on mothers' gender attitudes. Mothers who value children's STEM participation were expected to use more supportive strategies in play, for example by supporting the child in understanding the building instructions or by keeping the child on-task. Mothers who value STEM participation less were expected to use less supportive strategies in play, either by dominating the interaction to the exclusion of the child or by not assisting the child in understanding the task. Similar patterns were expected for mothers with more liberal gender attitudes, as well as for mothers with better spatial skill. Strategy coding is discussed further in the Video Coding section.

It was expected that children would show gender-typed preferences for the games, with girls would preferring *GoldieBlox* to *BobbyBlox* and boys preferring *BobbyBlox* to *GoldieBlox*. It was expected that girls would thus engage longer in independent free play if playing with *GoldieBlox* than with *BobbyBlox*. In contrast, boys were expected to engage longer in free play if playing with *BobbyBlox* than with *GoldieBlox*. Children were expected to show sex-typed modes of play with either set, with girls engaging in more reading and symbolic play than boys and with boys engaging in more construction. Gender-typed interest and mode of play were expected to vary with individual differences in attention to gender (e.g., GSF).

With respect to mechanical learning, because children's play was expected to be influenced by the gender-toy match, with children playing more if given the toy marketed to their gender group as compared to the other-gender toy, mechanical learning was expected to be influenced accordingly. Girls were thus expected to learn more from *GoldieBlox* than from *BobbyBlox*. Boys were expected to learn more from *BobbyBlox* than from *GoldieBlox*. This was also expected to translate to another task: Girls who played with *GoldieBlox* and boys who played with *BobbyBlox* were expected to be more successful on a transfer mechanical task than girls who played with *BobbyBlox* or boys who played with *GoldieBlox*. Children with stronger spatial skill were expected to perform better overall on the mechanical tasks than were children with weaker spatial skill. Mothers' play scaffolding strategies were expected to mediate the relationship between child variables (e.g., GSF, gender, spatial skill) and mechanical learning. This was predicted because mothers' strategies guide the dyadic interaction and toy use, and thus likely shape what children learn.

Finally, with respect to interest in subsequent STEM toys, it was expected that children would show gender-typed preferences, with girls preferring feminine-marketed STEM toys and boys preferring masculine-marketed STEM toys. Girls who played with *GoldieBlox* were also expected to be more interested in STEM toys overall than girls who played with *BobbyBlox*. This effect of condition was not anticipated for boys because most available STEM-relevant toys are already gender-typed as masculine and thus boys' interest was already anticipated to be quite high.

# Chapter 2

# METHOD

## Power Analysis

An apriori power analysis was used to determine sample size. Power for ANOVA and regression was estimated using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). The necessary sample size to detect a medium effect up to the two-way interaction (child gender X condition) with 80% power is 60 dyads. A medium effect size was chosen based on pilot research and on the recent meta-analytic finding that spatial interventions resulted in training gains with an average effect size of Hedge's $g = 0.47$ (Uttal et al., 2013). Other research with gendered toys and many of the same gender measures used in the present study found large effect sizes for girls' cognitive outcomes, equivalent to partial $\eta^2 = .31$ (Coyle & Fulcher, 2011). Even using the more conservative estimate for large effects of partial $\eta^2 = .14$, the 60 dyad sample is estimated to permit a regression design with condition, gender, an additional predictor, and a covariate (e.g., spatial skill), as well as all interactions, at 80% power.

## Participants

Families were recruited through letters sent home with children attending university or community preschools and summer camps in the Northeast U.S. Research announcements were posted on university and community notice boards. Participants were also recruited by email, using a laboratory-maintained database of families interested in being contacted for developmental science research participation. In total, 68 parents and their preschool-age children participated in the laboratory session.

Seven dyads were excluded from analysis. One dyad was excluded because the child was clinically diagnosed with a developmental disorder impacting his ability to understand and complete the laboratory tasks (Autism Spectrum Disorder). Four additional dyads included children who were siblings to children that had already participated in the study and were thus excluded. Finally, two dyads' data were not usable due to a high level of interference by younger siblings during the laboratory session.

Thus, the final sample included 61 mother-child dyads. Children (31 girls, 30 boys) ranged in age from 4 years, 0 months to 6 years, 11 months (*M* = 5 years, 2 months, *SD* = 10.34 months, median = 5 years, mode = 4 years, 3 months). Mothers ranged in age from 25 to 62 years old (*M* = 36.13 years, *SD* = 5.83, median = 35.50, mode = 32). Families were White (90%), African American (4%), Asian (4%), and Latino (2%). Families were majority middle- to upper-middle-income, with 26% reporting annual household income between $50,000-$80,000 and 38% between $80,000-120,000. Mothers were highly educated, with 43% holding a bachelor's degree and 39.3% holding a graduate degree. Most mothers were employed, with 28% working full-time and 36% working part-time. Fathers of participating children were similar in age, education, and employment to mothers. They ranged in age from 26 to 60 years old (*M* = 37.76 years, *SD* = 6.65, median = 36, mode = 33). Fifty-four percent of fathers held a bachelor's degree and 21.3% held a graduate degree. All fathers were employed, with 72% reporting working 40 or more hours per week and 21% reporting working 20-39 hours per week. Family characteristics are summarized in Table 1.

**Procedure**

The procedure is summarized in Table 2. Mother-child dyads were randomly assigned to play with *GoldieBlox* or with *BobbyBlox*, and were tested individually. Play was videotaped with three HD-PTZ wireless cameras: One camera was aimed at the table surface where mothers and children were playing, one camera was aimed at the child's face, and one camera was aimed at the mother's face. The three cameras were mounted in the corners of the testing space just below the ceiling. A condenser microphone was mounted on the ceiling to record audio. See Appendix B for the experimental script.

Mothers had approximately five minutes alone to familiarize themselves with the toy while a researcher surveyed the child about favorite toys and administered the water level task in a separate room. Mothers were told, "This is a new toy recently released to the public. We are interested in seeing how parents and children interact with it. Take a few minutes now to familiarize yourself with the toy and its pieces. I will bring your child back in about five minutes so you two can play together."

The mother and child were then given approximately fifteen minutes to play with the toy together. Mothers were not given specific instructions about how to play with the toy and if they asked for instructions, they were told to play however they might play typically at home. Following fifteen minutes of dyadic play, the child was then left to play individually for seven minutes. While children played alone, mothers were given the water level task and completed a demographic survey in the adjacent room. Mothers also made ratings about their reaction to *Goldie/BobbyBlox* and their interest in purchasing those and other STEM toys.

Following seven minutes of free play, an interval determined to be appropriate based on pilot testing, children were given a mechanical learning assessment with the toy, administered by the researcher. These problems, described further in the Mechanical and Spatial Measures section that follows, measure how well children learned the belt drive concept targeted in the toy. Children were then asked to complete an additional mechanical task to assess generalization and transfer of mechanical learning to another toy. This task is also described in the following section.

Finally, children were surveyed about how much they liked *Goldie/BobbyBlox*, how they might change the toy, and how interested they were in other STEM toys (described further in STEM Toy Interest section below). Parents and children were thanked and given a gift certificate for an ice cream cone and reimbursement for parking expense.

Approximately one week after the laboratory session, mothers received a link to a follow-up web survey. This survey assessed mothers' and children's gender attitudes and beliefs and was administered a week following the laboratory session due to time constraints as well as to avoid contamination by the laboratory session and vice versa. Mothers were instructed to complete their own surveys and to assist their child in navigating the child portion of the surveys. Mothers completed measures of feminism and sexism, division of labor, gender interests and attitudes, and gender socialization and beliefs about their children. The order of the surveys can be found in Table 2. Details about measurement are discussed in the following sections.

Children completed surveys about gender interests and attitudes and indicators of personal gender salience, including gender typicality, incidental memory for gender, and gendered affiliative preference. The survey order appears in Table 2. Survey details are discussed in the following sections. For the child surveys, mothers were instructed:

> Most of all, we are interested in what your child thinks. Thus, it is important that you not guide them in selecting a particular answer or correct their answers. You will need to read the instructions to your child and may need to assist in moving through the survey. Read the questions aloud in a neutral tone.
>
> Remember with these questions that there are no right or wrong answers. You might even disagree with your child! That is ok, we are simply interested in what your child thinks about a range of topics. Children develop ideas and interests over time. We are interested in the snapshot of what your child thinks right now.

For participating, families were entered into a raffle to receive a gift certificate to Toys 'R' Us. They received additional entries for each portion of the study that they completed (laboratory session, mother's web survey, child's web survey). Four children's names were drawn for the raffle. Families also received compensation for parking expenses for the laboratory session ($1-2).

**Toy Materials**

*GoldieBlox* is recently developed brand of engineering toys targeted at and marketed to girls. The first set was released to the public in January 2013. Now in January 2015, there are four themed kits, add-on packs of building materials, and a recently released action figure. Each kit aims to teach a simple machine or engineering skill and most pair a book with a building set. The present study used the first version of the first set released, *GoldieBlox and the Spinning Machine*. This set targets the belt-drive, a simple machine involving a belt around two or more wheels mounted on axles, resulting in motion. The basic story describes Goldie learning what makes the ballerina in a jewelry box move, and wanting to scale that machine up to spin, first her dog, and then all of her animal friends. Goldie is a blonde girl, aspiring to be an engineer. The set includes the book, a pegboard, axles, wheels, washers, blocks, a ribbon (belt), and five animal dolls (animals are depicted in Figure 1).

To test the role of marketing *GoldieBlox* explicitly as a toy for girls, I created a comparison toy marketed to boys (*BobbyBlox*). *BobbyBlox* was identical to *GoldieBlox* in form and function, using all the same pieces and the same story. The box design was adapted by hand and in Photoshop to be identical in design elements but traditionally masculine in content. The pink ribbon in the feminine game was replaced with a green ribbon. The story illustrations were adapted by hand and in Photoshop to remove feminine elements, recolor traditionally feminine colors, and to replace female pronouns with male pronouns. Images of Bobby were rated for masculinity by Introductory Psychology Subject Pool participants to verify that the illustrations were indeed masculine in nature (Figure 2). See a comparison of the two toys in Figure 3.

**Mechanical and Spatial Measures**

Spatial perception is one set of spatial skills where gender differences are typically found (Linn & Petersen, 1985; Wai, Cacchio, Putallaz, & Makel, 2010). To measure spatial perception, parents and children completed age-appropriate versions of the water-level task (Liben, 1995; Liben & Downs, 1986; Piaget & Inhelder, 1956). In this task, participants are shown a jar tipped at various angles and either select the corresponding drawing that shows where a water line would be in that jar (for children; e.g., “Which one shows where the water would be when the jar is in this position?”; Liben & Downs, 1986) or produce the line to show where the water would be (for adults; e.g., “Please draw a line to show where the water would be in that bottle while it is being held in the position shown"; Liben, 1995). The 6-item adult task is scored by measuring the number of degrees each line is off of horizontal and summing the number of items correct within 5 degrees (most accurate) and within 10 degrees (allows slightly more error). The 6-item child task is scored by summing the number of instances where the child correctly chooses the jar with the horizontal water line. Although this measure has been used with young children before (Liben & Downs, 1986), use with preschool children has not been well-established. Reliability was low for children in the present study (alpha = .50). Spatial perception is a developmental process posited to occur during this period (Piaget & Inhelder, 1956) thus reliability was also calculated split by age. Children younger than 60 months were most inconsistent in their responses (alpha = .25) whereas children 60 months and older were relatively more consistent (alpha = .60).

Mechanical learning was measured at the conclusion of play with *Goldie/BobbyBlox* through a series of problems designed to test understanding of the belt-drive concept targeted in the toy. The measure was comprised of a series of five problems given in roughly ascending order of difficulty, determined based on pilot testing. Children were asked to set up a belt drive, first with two wheels (for the first two problems) and then with three wheels (for the final three problems). They were asked to select characters to sit on the wheels for ease of visualizing the direction the wheels turned. The belt drive problems involved building the drive with two or three wheels so that the characters turned in the same or in opposite directions, as directed in the problem. Mechanical learning was scored as the number of problems the child solved correctly, ranging from 0 to 5. Additionally, children's choice of characters to sit on the wheels used in constructions to solve the problems (3 animals) was scored for gender-typing. The five potential characters were rated in pilot-testing for degree of masculinity or femininity (see Table 3) and this continuum was used to create a score for each child averaging their choice of the first two animals selected for the mechanical problems.  See Appendix C for the full mechanical learning measure.

To measure the extent to which mechanical learning generalized to a mechanical task outside of the *Goldie/BobbyBlox* toy, children completed two additional transfer tasks. For both tasks, children were given an image of a structure built with Tinker Toys that utilized wheels and axles similarly to *Goldie/BobbyBlox*. Children were instructed to build the structure in each image. For the first task (*Easy Transfer Task*), children assembled a simple three-wheeled machine. Because nearly all children completed the Easy Transfer Task correctly, this task was scored for time to complete building, with faster building generally reflecting more fluency with construction. For the second task (*Difficult Transfer Task*), children assembled a larger and more complex four-wheeled vehicle with numerous asymmetries. Constructions were scored for accuracy (number of errors) and completion at two time points: After seven minutes of building and when children stopped building -- because they finished building, they were unable to finish building, or they ran out of time in the laboratory session. Specific construction substitutions were permitted (i.e., a longer axle could be substituted for a shorter axle because they served the same mechanical function). Scores ranged from 1 (unfinished with 2+ errors) to 4 (finished correctly). The transfer task measures can be found in Appendix D (*Easy Task*) and Appendix E (*Difficult Task*). Descriptive statistics for all spatial and mechanical measures can be found in Table 4.

**STEM Toy Interest**

To assess the degree to which mothers and children liked *Goldie/BobbyBlox* and to measure their interest in other STEM toys, mothers and children were separately queried about their interests. Mothers were asked to rate how much they enjoyed playing with *Goldie/BobbyBlox* with their child, how much educational value they toy had, and what they would change to improve the toy (open-ended). Mothers then rated their interest in purchasing *GoldieBlox*, *BobbyBlox*, and a series of 15 STEM toys that were marketed as masculine (5 items), feminine (5 items), or neutral (5 items), presented in a single, randomly determined order. Toys mothers rated included super hero LEGOs (masculine), equestrian LEGO Friends (feminine), and tangrams (neutral). Ratings were made on a 4-point Likert scale from “Not at all interested” to “Very interested” and reliability was adequate (alpha = .71). Descriptive statistics for mothers’ interests appear in Table 5.

Children responded to a similar series of questions. Children made ratings using a 3-point pictorial scale, “not at all,” “some,” or “a lot,” accompanied by graphic representations of glasses that were completely empty, half-full, or completely full. Children rated how much they enjoyed playing with Goldie/Bobby Blox and how much they would like to have the toy at home versus at school. They were also asked how they would change the toy to improve it (open-ended).

Children then completed a perceived typicality scale similar to that by Coyle & Liben (2013). Children were told to think about the character in the game they had just played (either Goldie or Bobby). They were then asked to compare themselves to the character, in a series of eight questions. Sample items were "Think about how Goldie [Bobby] looks. How much are you like that?" and "Think about what Goldie [Bobby] is good at. How much are you like that?" Children used the same 3-point pictorial scale for their answers. Finally, children rated the same 15 STEM toys that mothers rated, responding to the question "How much would you like to play with this toy?" Again, children used the 3-point pictorial scale. Reliability was good (alpha = .71). Descriptive statistics for children's interests appear in Table 6. STEM toys used for mother and child ratings appear in Appendix F. The perceived similarity to character scale can be found in Appendix G.

**Division of Paid and Unpaid Labor**

Parent's gender attitudes, beliefs, and behavior were measured in a variety of ways, discussed in this and the following subsections. Gendered behavior of particular interest was division of labor within and outside of the home. To assess families division of household labor, mothers completed the 44-item Who Does What survey (Cowan & Cowan, 1990). This three-subscale measure asks how household labor, decision-making, and childcare are divided between partners. A sample item from the household labor subscale is "Cleaning up after meals," to which participants rate who is responsible. A sample item from the decision-making subscale is "Deciding about major expenses: house, car, furniture." A sample item from the childcare subscale is "Cleaning or bathing our child." Ratings are made on a 9-point Likert scale from "My partner does it all" (1) to "We both do this about equally" (5) to "I do it all" (9). It has been shown to have adequate reliability in previous research and in scale development (Cowan &

Cowan, 1990), and all three subscales had good reliability in the present study (alphas = .82, .87, .93 for household labor, decision-making, and childcare subscales respectively).

Gendered behavior outside of the home was indexed by occupation and work behaviors. In the demographic survey administered during the laboratory session, mothers reported their own and their partner's education, occupation, and hours employed outside the home. For mothers who were not employed and stayed home full time, this was considered their occupation, with no hours employed outside the home. Full-time childcare is a highly gender-traditional vocation for mothers. Reported occupations were rated by Introductory Psychology Subject Pool participants for gender traditionality and for perceived occupational prestige. Gender traditionality ratings were made on a 5-point Likert scale from “Very feminine” to “Neither feminine nor masculine” to “Very masculine.” Occupational prestige ratings were made on a 4-point Likert scale from “Not at all prestigious” to “Extremely prestigious” Work hours were assessed categorically as no work at all, part time work, and full time work. Gender traditionality, occupational prestige, and employment hours were standardized and combined to create a composite score that reflects gender traditionality of mothers' and fathers' work. To permit mothers' and fathers' scores to be comparable, scales were left in the direction in which they were assessed. Thus for mothers, high scores on the work traditionality composite indicate greater gender nontraditionality, or a job that is less traditionally feminine, higher prestige, and involves more hours employed outside the home, whereas low or negative scores on the composite indicate a job that is highly traditionally feminine, low prestige, and involves fewer hours employed outside the home. For fathers, a high score reflects a job that is highly traditional, high prestige, and involves many hours employed outside the home. Descriptive statistics for families division of paid and unpaid labor appear in Table 7.

**Mothers' Gender Attitudes and Beliefs**

**Gender interests.** To assess the degree to which mothers held culturally feminine or masculine interests and traits, they were given the activity and trait subscales of the personal measure (PM) of the Occupations, Activities, and Traits (OAT) sex-typing scale (Liben & Bigler, 2002). The OAT-PM contains 10 feminine, 10 masculine, and 5 neutral filler items given in a single randomly determined order. The activities subscale asks how often one engages in a series of activities whereas the traits subscale asks to what degree a trait is self-descriptive. Sample items from the OAT-PM are "bake cookies" (feminine) and "play darts" (masculine). Sample items from the OAT-PM are "affectionate" (feminine) and "aggressive" (masculine). Ratings were made on a 4-point Likert scale from "Never" to "Often or very often" for the OAT-PM and from "Not at all like me" to "Very much like me" for the OAT-PM. The OAT-PM has been shown to have good reliability in previous research and in scale development (Liben & Bigler, 2002). However, they had low reliability in the present study (alphas = .44, .54, .65, .63 for feminine activities, masculine activities, feminine traits, and masculine traits respectively). I return to the OAT-PM reliability in the Results section. Descriptive statistics can be found in Table 7.

**Gender attitudes.** To assess the degree to which mothers' attitudes toward others were sex-typed versus more flexible, mothers completed the activity subscale of the attitude measure (AM) of the OAT (Liben & Bigler, 2002). This measure is similar in item content to the OAT-PM, including 10 feminine, 10 masculine, and 10 neutral activities. Sample items included "read romance novels" (feminine) and "build with tools" (masculine). Mothers rated on a 5-point scale whether each activity should be done by "Only men," "Mostly men, some women," "Both men and women," "Mostly women, some men," or "Only women." Scores were then recoded to reflect flexible (“Both men and women”) versus stereotyped responses. The proportion of answers where the flexible choice was selected indicated mothers’ gender attitude flexibility. This scale has shown adequate reliability in previous use and in scale development (Liben & Bigler, 2002) and was reliable in the present study (alpha = .86).

Mother’s feminist attitudes were measured using Becker and Wagner’s (2009) 5-item Collective Action subscale. A sample item from this subscale includes “I make a conscious attempt to use non-sexist language.” Ratings were made on a 6-point Likert scale from “Disagree strongly” to “Agree strongly.” It has shown good reliability in previous research and in scale development (Becker & Wagner, 2009), and was highly reliable in the present study as well (alpha = .85).

To measure sexism, mothers completed the 8-item Modern Sexism scale (Swim, Aikin, Hall, & Hunter, 1995). A sample item from this scale is “Discrimination against women is no longer a problem in the United States.” Ratings were made on a 5-point Likert scale from “Strongly disagree” to “Strongly agree.” It has shown adequate reliability in previous research and in scale development (Swim et al., 1995), and was highly reliable in the present study (alpha = .80). Descriptive statistics for mothers’ gender attitude measures can be found in Table 7.

**Gender socialization.** Mothers’ gender attitudes may be reflected in their active socialization of gender through parenting. The 28-item Gender Socialization Scale (Blakemore & Hill, 2008) assesses how parents feel about their son or daughter engaging in various behaviors gender-typed behaviors, as well as how parents encourage or discourage their children in different ways. The six subscales are: Feminine activities, masculine activities, helping around the house, education for future career, education for future family, and discouraging gender nonconformity. For the first three subscales, mothers rated how they would feel if their child engaged in a particular behavior or activity, using a 7-point Likert scale from “Very negative” to “Very positive.” Sample items include “playing with a toy nurse kit” (feminine activity), “playing with toy cars” (masculine activity), and “setting the table” (helping around the house). For the remaining subscales, mothers rated their agreement with statements, using a similar 7-point scale from “Disagree strongly” to “Agree strongly.” Sample items include “I would encourage my child to go to college” (education for future career), “I would want my child’s education to prepare him/her for child rearing” (education for future family), and “I would discourage my child from acting like a member of the opposite sex” (discouraging gender nonconformity). This scale has shown good reliability in previous research and in scale development (Blakemore & Hill, 2008), and showed generally good reliability in the present study (alpha = .85, .74, .81, .31, .83, .64 for feminine activities, masculine activities, helping around the house, education for future career, education for future family, and discouraging nonconformity respectively). The low alpha for the education for future career subscale is likely due to the restricted range of responses to those questions. Nearly all mothers said they “Agree strongly” that their child’s education should prepare them to earn and living and that their child should go to college. Descriptive statistics for mothers’ gender socialization, split by child gender, can be found in Table 8.

***Expectancy-value.*** One way that gender socialization and parents’ gender attitudes can manifest is through expectations about academic performance and future occupations. Mothers completed a parent-child expectancy value scale about their child’s academic performance (Simpkins et al., 2012). The first subscale, Perceptions of Child’s Ability (3 items), assessed how talented mothers believed their son or daughter was in four domains: Sports, music, math, and reading. Ratings were made on a 7-point Likert scale from “Not at all good” to “Very good.”

Reliability was high in previous work (Simpkins et al., 2012) and in the present study (alpha's ranged between .78 and .88). The second subscale, Perceptions of Importance, asked how important performance was in each of the four domains. Ratings were made on a 7-point scale from "Not at all important to "Very important." Each question was a single item considered independently. Finally, mothers rated their own efficacy to assist their child in each of the four activities, on a 7-point scale from "Much less confident than other activities" to "Much more confident than other activities." Again, each item was a single question considered independently.

Mothers also rated how positively or negatively they would feel about their child holding various jobs. Jobs were taken from the COAT-PM, the Children: Occupations, Activities, and Traits occupations subscale of the personal measure. Similar to the OAT-PM subscales described previously, this subscale includes 10 feminine jobs, 10 masculine jobs, and 5 neutral jobs. Mothers made ratings using the 7-point Likert scale from "Very negative" to "Very positive" that appeared in the Gender Socialization scale. Reliability for these scales were high (alphas = .91 for feminine jobs and .72 for masculine jobs). Descriptive statistics for expectancy-value measures can be found, split by child gender, in Table 8.

**Children's Gender Attitudes and Beliefs**

**Gender interests.** To assess the degree to which children held culturally feminine or masculine interests, children completed the occupation and activity subscales of the personal measure of the preschool version of the Occupations Activities and Traits scale (POAT; Liben, Bigler, Shechner, & Arthur, 2006). This scale is similar to the OAT-PM in structure and intention but is adapted to be age-appropriate for preschoolers: It is shorter, uses graphic representations of the activities and traits, and uses a shorter rating scale. Children used the same

3-point Likert water glasses scale used in the Goldie/Bobby Character Typicality measure. The POAT-PM contains 6 feminine, 6 masculine, and 2 neutral filler items given in a single, randomly determined order. The occupations subscale asks how interested the child is in engaging in various occupations as an adult whereas the activity subscale asks how much the child likes to engage in various play activities. Sample items from the P<u>O</u>AT-PM are “ballet dancer” (feminine) and “police officer” (masculine). Sample items from the PO<u>A</u>T-PM are “dress-up clothes” (feminine) and “robots” (masculine). The POAT-PM has been shown to have good reliability in previous research and in scale development (Liben et al., 2006), and had good reliability in the present study as well (alphas = .73, .80, .90, .72 for P<u>O</u>AT-PM feminine, masculine, PO<u>A</u>T-PM feminine, and masculine, respectively).

Children’s gendered interests were also gleaned from their self-reported favorite toys. Children responded to the question “What is your favorite toy to play with at home?” Introductory Psychology Subject Pool participants rated children’s reported favorite toys for femininity or masculinity, on a 5-point Likert scale from “Very feminine” to “Neither feminine nor masculine” to “Very masculine.” Descriptive statistics for children’s gendered interests are reported, split by gender, in Table 9.

***Gender-typed appearance.*** One final indicator of children’s gender interests was children’s clothing. To measure the degree to which children adhered to rigidly gender-typed clothing, children’s clothes were coded following the procedures developed by Halim and colleagues (Halim, Ruble, Tamis-LeMonda, & Shrout, 2013). Still photographs were taken from the videos of the laboratory session, showing the child’s full body, including shoes, hairstyle, and any design or pattern on clothing items. The photographs were coded for the presence or absence of a set of gender-typed elements. For girls, codes included whether she was wearing a dress, the presence of feminine colors and patterns, the presence of current trendy items, feminine hairstyle, and jewelry. For boys, codes included the presence of masculine colors, masculine logos or themes, masculine formalwear, masculine fabrics, and sports-related apparel. Scores were the sum of the number of gender-typed elements present. For girls, the range was 0 to 8 while for boys, the range was 0 to 5. Because the ranges differed, scores were standardized within gender to create comparable scores to represent the degree to which children were gender-typed in appearance. Descriptive statistics are presented in Table 9.

**Gender attitudes.** To assess the degree to which boys' and girls' held sex-typed versus flexible attitudes towards others, children completed the occupation and activity subscale of the attitude measure of the POAT (Liben et al., 2006). This was parallel to OAT-AM in general structure and content and similar to the POAT-PM in age-appropriate adaptation. The POAT-AM included 6 feminine, 6 masculine, and 2 neutral items. Sample items from the PO̲AT-AM include "nurse" (feminine) and "truck driver" (masculine). Sample items from the POA̲T-AM include "doll houses" (feminine) and "bulldozers" (masculine). Children indicated on a 3-point scale whether each occupation or activity should be done by "Only men" (depicted by two male bathroom icons), "Both men and women" (depicted by one male and one female bathroom icon), or "Only women" (depicted by two female bathroom icons). As with the OAT-AM, scores were then recoded to reflect flexible versus stereotyped responses. The proportion of answers where the flexible choice was selected indicated children's gender attitude flexibility. The POAT-AM has shown good reliability in previous use and in scale development (Liben et al., 2006) and was reliable in the present study (alphas = .73 and .78 for PO̲AT-AM and POA̲T-AM respectively).

***Gender salience filter.*** In previous work (Coyle & Liben, 2013), GSF has been measured as the composite of three components: (a) gendered affiliation, (b) incidental memory for gender, and (c) gender typicality. The same measures were used in the present study, delivered in the child portion of the web survey.

Serbin and Sprafkin's (1986) measure of gendered affiliation was used to measure the degree to which gender was a salient category for children for affiliative preference. Children were first shown five pairs of photos, each with a man and a woman with neutral expressions and no props, and asked with whom they would prefer to play. Children were given a second set of faces based on their choices in the first set. Both second sets again contained five pairs of faces, each with a man and woman. If the child selected mostly men in set one, the pairs in the second set showed men with blank expressions and women with smiling expressions and holding an interesting game or toy. If the child selected mostly women in set one, the second deck showed the inverse parings, that is, men with blank expressions and no props, and women smiling and holding a toy. The affiliation score was the proportion of selections in set two that matched the predominant gender preference in deck one. A high score indicates that the child consistently chose play partners based on preferred gender, even when that meant rejecting a friendlier play partner of the other gender. In contrast, a low score indicates that the child does not consistently use gender as a basis for selecting play partners and perhaps instead uses another criteria (e.g., presence of toys). This scale has good reliability in scale development and previous use (Serbin & Sprafkin, 1986) and was adequate in the present study as well (alphas = .74 for girls and .65 for boys).

Liben and Hilliard's (2010) measure of incidental memory for gender in classroom contexts was used here to measure incidental memory for gender more generally. The measure assesses the degree to which children incidentally encode information about gender while making ratings about 10 gender-neutral toys (using the 3-point water glass scale described previously). First, children were shown 10 images of a toy being played with by a girl (5 toys) or a boy (5 toys) and rated how much they would like to play with the toy. The gender-toy pairings were counterbalanced across participants and toys were presented in a single, randomly determined order. After a delay, children were shown images of the 10 toys, now depicting only the toy and not the child. For each toy, children were asked whether the child originally depicted playing with the toy was a girl or a boy. Incidental memory for gender was therefore the proportion of items for which gender was correctly recalled. This measure had good reliability in previous work (Coyle & Liben, 2013; Liben & Hilliard, 2010) and was highly reliable in the present study (alpha = .81).

To assess children's self-perceived gender typicality, I used a measure Coyle and Liben (2013) adapted for preschool children from Patterson's (2012) measure for school-age children. Children rated their similarity to girls, boys, and children as a group ("kids") in terms of appearance, interest, competencies, and behaviors on the 3-point water glass scale. This was similar to the scale used for Goldie/Bobby Blox Character Typicality, described previously. Sample items include "Think about the things that most girls do. How much are you like that?" and "Think about the things that boys are good at. How much are you like that?" Responses to items about similarity to girls and boys were averaged separately to yield scores between 1 and 3 indicating self-perceived similarity to girls and to boys. Own-sex typicality, or similarity to one's own group (girls, for girl participants and boys, for boy participants) was used in the composite GSF score discussed next. This scale has shown good reliability in previous use (Coyle & Liben, 2013) and was very reliable in the present study (alphas = .95 and .87 for similarity to girls and to boys respectively).

Previous work (Coyle & Liben, 2013) has defined GSF as the summed standard scores for gendered affiliation, gender memory, and own-sex affiliation. Coyle and Liben (2013) validated this approach through factor analysis and thus, their approach is the one used here as well. The three GSF components were standardized and summed to create the GSF composite. Descriptive statistics for the three components and the composite score are in Table 9, split by child gender.

# Chapter 3

# RESULTS

## Video Coding

Videos of laboratory session were coded using INTERACT observational coding software (Mangold, 2013). Coding was done independently by pairs of research assistants (20% overlap). Interval coding was selected for use in the current study. Interval coding breaks each video into intervals of equal length and the coding system is applied to each interval. Reliability is calculated for coder match across intervals rather than for match across intervals and accuracy of time, as can become an issue with coder-determined duration coding. Intervals used in the current study vary across coding systems and range from 5 to 10 seconds depending on the system. The length of the interval was determined based on behavior frequency and length, co-occurrence of behaviors, video length, and coder feedback during system development. Interval lengths were optimized to capture as much variability in behavior as possible and to limit the occurrence of multiple codes per interval, especially for systems where only one code per interval was permitted.

The mother familiarization videos were coded in 10-second intervals with one code per interval (Kappa = .97). Within an interval, the majority behavior was coded for the interval. See Table 10 for a description of the mother familiarization coding system. To describe the manner in which the dyad played, dyadic play behaviors were coded in 8-second intervals with one code per interval (Kappa = .78). As with the mother familiarization, the majority behavior for the interval was coded. See Table 11 for a description of the codes used in the dyadic play behavior coding system. Mother scaffolding during dyadic play was coded in 10-second intervals (Kappa

= .74). In this coding system, multiple codes were permitted per interval. See Table 12 for a description of the codes in this system. Finally, child play was coded in 5-second intervals with a single code per interval (Kappa = .91). Within each interval, the majority behavior was coded for the interval. Table 13 describes the child play coding system.

Data were exported from INTERACT into SPSS as relative frequencies. The use of relative frequencies permits the proportional comparison of code frequency across videos of different length and across codes that co-occur in a single interval (as with mother scaffolding codes). Thus, the relative frequency of a code is the frequency of that code for a dyad, controlling for the frequency of other codes for the same dyad. Relative frequencies of codes in each system are discussed in the analyses that follow.

**Preliminary Analysis**

To determine the need for covariates, I first examined correlations among demographic variables (child age, parent education, socioeconomic status, and race) and key study variables. There were significant correlations between child age and performance on the mechanical learning post-test and transfer tasks ($r$ = .33, -.49, and .43, $p$s = .009, <.001, .001, respectively, for mechanical learning, time to complete the easy task, and performance on the hard task). Children's age was thus included as a covariate in subsequent analyses related to mechanical learning. There was also a significant correlation between children's age and how interested they were in neutral STEM toys ($r$ = .38, $p$ = .002). Age was thus also included as a covariate in analyses related to STEM toy interest. Correlation matrices among key study variables for children (split by gender) are presented in Table 14 (*GoldieBlox*) and Table 15 (*BobbyBlox*). Correlations among key study variables for mothers (split by child gender) are presented in Table 16 (*GoldieBlox*) and Table 17 (*BobbyBlox*).

To assess the comparability between the experimental conditions formed by random assignment, I compared the two groups' performance on the water-level task, which was assessed very shortly after arriving for the laboratory session. This analysis revealed no significant difference between the groups in spatial ability, $t(59) = 0.213$, $p = .832$. Because age was correlated with mechanical performance and considered a covariate for later analysis, I also compared the two groups on age to ensure there was no difference in mean age across the conditions. No age difference was found between the groups, $t(59) = 1.58$, $p = .119$.

Because mothers participated in the laboratory session and fathers did not, mothers were compared to fathers on demographic characteristics to establish where important gender differences might lie. Mothers were the sole reporter of division of household labor, decision-making, and childcare. Scores on this measure could range from 1, father does all, to 9, mother does all. On average, mothers reported being responsible for a slight majority of household labor ($M = 6.05$, $SD = 1.13$, median = 5.62, mode = 5.54, min = 5.48, max = 9.00). Mothers reported being responsible for a somewhat larger proportion of childcare ($M = 6.89$, $SD = 1.13$, median = 6.95, mode = 5.63, min = 3.00, max = 9.00). Mothers generally reported equal decision-making between partners ($M = 5.71$, $SD = 1.06$, median = 5.40, mode = 5.00, min = 4.17, max = 9.00). To compare the gender traditionality and prestige of mothers' and fathers' occupations, I conducted a series of paired *t*-tests with Sidak correction to the alpha level. Mothers held more feminine occupations ($M = 2.20$, $SD = 0.92$) while fathers held more masculine occupations ($M = 3.31$, $SD = 0.56$), $t(59) = 8.54$, $p < .001$. Mothers also held less prestigious occupations ($M = 1.96$, $SD = 0.68$) than did fathers ($M = 2.61$, $SD = 0.52$), $t(59) = 6.13$, $p < .001$. Maternal and paternal occupational traditionality (the sum of the standardized scores for occupational gender traditionality, occupational prestige, and hours employed) are used later in analysis to test the role of parents' gender attitudes and behaviors. For mothers in particular, work traditionality was used in place of the OAT-PM measures of gender personal interests due to their low reliability in the present study (see Method section).

Descriptive statistics for children's spatial and mechanical outcomes (split by gender) can be found in Table 4. Descriptive statistics for children's game interest outcome measures can be found in Table 5, split by gender and condition. Descriptive statistics for children's individual difference measures are reported in Table 9. Descriptive statistics for mothers' spatial measures are reported in Table 4. Descriptive statistics for mothers' individual difference measures can be found in Tables 1 and 7. Mothers' beliefs about their children are described in Table 8. Histograms depicting sample distribution on key demographic and study variables are presented in Appendices H-L. Appendix H depicts distributions for family characteristics. Distributions of individual difference variables are in Appendix I (mothers) and J (children). Distributions of outcome variables are in Appendix K (mothers) and L (children).

**Overview of Analyses**

As explained in the introductory chapter, I expected that children would be sex-typed in their interest in the toy and that this would shape their learning. I further expected that mothers would differentially scaffold play between sons and daughters. In the following subsections, I report on analyses to test the specific hypotheses discussed earlier, with respect to mothers' and children's play with, interest in, and learning from *GoldieBlox* versus *BobbyBlox*. I first present the analyses testing mothers' and children's play with the toy in (a) mother's familiarization, (b) dyadic play, and (c) children's free play. I then present analyses testing children's and mother's interest in *Goldie*/*BobbyBlox*. Finally, I present learning outcomes specific to STEM, including (a) mechanical learning, (b) transfer of mechanical learning to another task, and (c) interest in other STEM toys. Most hypothesis testing is done using mixed model analysis of variance (ANOVA). All follow-up analyses used to decompose higher-order interactions report Sidak-adjusted *p*-values to correct for multiple comparisons. When the assumption of sphericity was violated, Mauchley's test is reported and the appropriately corrected degrees of freedom and *p*-values are reported (Greenhouse-Geisser when ε was less than .75 and Huynd-Feldt when ε was equal to or greater than .75; Kirk, 2013).

**Mother Familiarization**

Because the toys were designed to be differentially marketed as being "for girls" versus "for boys," mothers were expected to use them differently beginning even with familiarization. Mothers given *GoldieBlox* were expected to spend more time reading than mothers given *BobbyBlox*, while mothers given *BobbyBlox* were expected to spend more time building than mothers given *GoldieBlox*. To compare relative frequencies of reading versus building during mother's familiarization with the toy, I conducted a 2(familiarization activity: reading vs. building) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) mixed model ANOVA in which familiarization activity was a within-subjects variable and condition and child gender were between-subjects variables. The analysis showed a main effect of familiarization activity, such that mothers engaged in more reading ($M$ = .63, $SD$ = .22) than building ($M$ = .06, $SD$ = .09) overall, $F(1,53) = 299.77$, $p < .001$, $\eta^2_p = .85$. There was also significant interaction between familiarization type and condition, $F(1,53) = 3.90$, $p = .050$, $\eta^2_p = .07$. Post-hoc pairwise comparisons showed that mothers engaged in more building when playing with *BobbyBlox* than with *GoldieBlox* ($M$ = .07 vs .05, $SD$ = .10 and .08; Sidak-adjusted $p = .035$), in partial support of the hypothesis. Remaining main effects and interactions were not significant. The proportion of mothers who engaged in any building during toy familiarization with *GoldieBlox* versus *BobbyBlox* is depicted in Figure 4.

**Dyadic Play Behaviors**

To understand how mother-child dyads allocated their play time to various activities, I conducted a 5(play behavior: book building vs. other building vs. reading vs. examining pieces vs. other activity) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) mixed model ANOVA in which play behavior was a within-subjects variable and condition and child gender were between-subjects variables. Mauchley's test of sphericity indicated that the assumption of sphericity had been violated, $\chi^2(9) = 49.95$, $p < .001$, $\varepsilon = .71$, thus the Greenhouse-Geisser correction was applied. There was a main effect of play behaviors, $F(2.83,161.34) = 63.61$, $p < .001$, $\eta^2_p = .71$. Post-hoc pairwise comparisons indicated that building along with the book and reading the story were the most common activities (Sidak-adjusted *p*s between those activities and all others were $<.001$, $p$ between reading and building with the book was .125). There were no significant differences in relative frequency of other play behaviors (Sidak-adjusted $p = .196$ between other building and examining pieces, .392 between other building and other play, and 1.000 between examining the pieces and other play). Remaining main effects and interactions were not significant. Proportion of time allocated to various play behaviors is depicted in Figure 5.

**Maternal Scaffolding**

I predicted that mothers would use more supportive scaffolding strategies with sons than with daughters, and that moms who especially valued STEM or feminism would also use more supportive strategies. First, to understand what strategies could be considered supportive of learning, I examined the correlations among maternal scaffolding strategies, children's mechanical learning, and game completion (see Table 18). Mechanical learning and game completion were independent outcomes and were not correlated ($r$ = -.05, $p$ = .732), perhaps indicating they represent alternative goals. Reading the book was the only strategy to significantly positively correlate with mechanical learning ($r$ = .30, $p$ = .020). Game completion, or how far the dyad got in the game during play independent of mechanical learning, significantly positively correlated with demonstrating building ($r$ = .37, $p$ = .004), and marginally positively correlated with correcting mistakes ($r$ = .23, $p$ = .080). Game completion negatively correlated with reading the book ($r$ = -.30, $p$ = .019).

To further test the strength of the scaffolding strategies as predictors of (a) mechanical learning and (b) game completion, I ran two forward regressions including all five scaffolding strategies. For mechanical learning, the only predictor to be added to the model was reading the book, $F(1,59) = 5.73$, $p$ = .020, $R^2$ = .09, $\beta$ = 0.30, $p$ = .020. For game completion, two predictors were added to the final model, $F(2,56) = 6.93$, $p$ = .002, $R^2$ = .20. Both demonstrating building ($\beta$ = 0.38, $p$ = .002) and correcting mistakes significantly predicted game completion ($\beta$ = 0.24, $p$ = .046). Reading the book was therefore used in future analysis pertaining to scaffolding strategy use supportive of mechanical learning, in addition to other strategies where appropriate.

To determine how and when mothers were employing various scaffolding strategies, I conducted a 5(scaffolding strategy: demonstrating vs. instructing vs. correcting mistake vs. labeling vs. reading book) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) mixed model ANOVA in which scaffolding strategy was a within-subjects variable and condition and child gender were between-subjects variables. Mauchley's test of sphericity indicated that the assumption of sphericity had been violated, $\chi^2(9) = 134.73$, $p$ < .001, $\varepsilon$ = .56, thus the Greenhouse-Geisser correction was applied. The analysis showed a main effect of scaffolding type, $F(2.22,126.73) = 80.39$, $p < .001$, $\eta^2_p = .59$. Post-hoc pairwise comparisons revealed that mothers were instructing and reading at the highest relative frequency (Sidak-adjusted *p* between instructing and reading was .958, *p*s between these codes and correcting or labeling were < .001, and < .01 for demonstrating). Mothers demonstrated less often than they instructed or read, and corrected mistakes and labeled pieces and constructions least often (Sidak-adjusted *p* between labeling and correcting mistakes was .676). The proportion of time mothers' allocated to each strategy is depicted in Figure 6, split by child gender Analysis also revealed a significant scaffolding type by gender interaction, $F(2.22,126.73) = 5.08$, $p = .006$, $\eta^2_p = .03$. Contrary to expectations, mothers utilized supportive scaffolding (i.e., reading) more often with daughters than with sons ($M = .40$ vs. .30, $SD = .13$ and .14, Sidak-adjusted $p = .009$). Mothers more often demonstrated building with sons than with daughters ($M = .27$ vs. .16, $SD = .18$ and .12, $p = .011$), a strategy correlated with game completion but not with learning. This interaction is depicted in Figure 7. There were no other significant main effects or interactions.

**STEM importance.** Mothers who especially valued children's performance in STEM were expected to employ strategies more supportive of mechanical learning. Given the earlier findings that (a) reading was the only strategy supportive of mechanical learning, (b) correcting mistakes and demonstrating building were supportive of game completion, and (c) mothers used reading with daughters but demonstrated building with sons, subsequent analyses of scaffolding strategy use include only these three strategies. To examine the role of STEM importance in strategy use, I conducted a 3(scaffolding strategy: demonstrating vs. reading vs. correcting mistakes) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(math importance: high vs. low) mixed model ANOVA in which scaffolding strategy was a within-subjects variable and condition, gender, and math importance were between subjects variables. Mothers who completed the expectancy value scale at post-test ($n$ = 44) were split into high and low math importance groups using a median split (median = 6.00). Mauchley's test of sphericity indicated that the assumption of sphericity had been violated, $\chi^2(2) = 14.08$, $p = .001$, $\varepsilon = .75$, thus the Huynh-Feldt correction was applied. As with the full 5 x 2 x 2 mixed model ANOVA, this analysis also showed a main effect of scaffolding, $F(1.86,66.90) = 39.94$, $p < .001$, $\eta^2_p = .53$ whereby mothers were most often using reading as a scaffolding strategy, demonstrating building less often, and correcting mistakes least often. There was a marginally significant interaction between math importance and scaffolding, $F(1.86,66.90) = 2.66$, $p = .082$, $\eta^2_p = .07$. Post-hoc tests showed that mothers who valued math less employed reading as a strategy even more often than did mothers who valued math more ($M$ = .41 vs. .31, $SD$ = .08 vs. .10, Sidak-adjusted $p = .065$). This interaction is depicted in Figure 8. Analysis did not show a main effect of math importance nor any additional interactions.

**Feminism.** Mothers with strong feminist values were also expected to employ more supportive strategies during dyadic play. I again conducted a 3(scaffolding strategy: demonstrating vs. reading vs. correcting mistakes) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(feminism: high vs. low) mixed model ANOVA in which scaffolding strategy was a within-subjects variable and condition, gender, and feminism were between subjects variables. As with math importance, mothers were split into high- and low-feminism groups using a median split (median = 2.20). As with math importance, dyads included in this analysis were those where the mother had completed the feminism scale in the post-test survey ($n$ = 44). Mauchley's test of sphericity indicated that the assumption of sphericity had been violated, $\chi^2(2) = 10.96$, $p = .004$, $\varepsilon = .79$, thus the Huynh-Feldt correction was applied. In parallel with the larger model, there was a main effect of scaffolding type, such that mothers were reading more often than demonstrating or correcting mistakes, $F(1.96,70.43) = 57.06$, $p < .001$, $\eta^2_p = .61$, and there was a scaffolding by child gender interaction such that mothers were reading more often with daughters and demonstrating building more often to sons, $F(1.96,70.43) = 5.45$, $p = .007$, $\eta^2_p = .13$. Analysis showed no main effect of feminism, however there was a significant four-way interaction among scaffolding, child gender, condition, and feminism, $F(1.96,70.43) = 6.66$, $p = .002$, , $\eta^2_p = .16$. Post-hoc comparisons showed that for dyads playing with *GoldieBlox*, highly feminist mothers demonstrated building more to sons than to daughters ($M$ = .35 vs. .16, $SD$ = .19 and .06, Sidak-adjusted $p = .019$) and read to daughters more than to sons ($M$ = .42 vs. .20, $SD$ = .14 and .12, $p = .018$). For dyads playing with *BobbyBlox*, low feminist mothers showed the same pattern, demonstrating building more often to sons than to daughters ($M$ = .28 vs. .12, $SD$ = .12 and .14, Sidak-adjusted $p = .035$) and reading to daughters more than to sons ($M$ = .41 vs. .25, $SD$ = .14 and .18, $p = .059$). This interaction is depicted in Figure 9. There were no other main effects or interactions.

**Mothers' work traditionality.** To permit examining the role of individual differences in mothers' gendered behaviors and attitudes with the full sample ($N = 61$), I also examined used of scaffolding strategies in relation to mothers' work traditionality. Gender-nontraditionally working mothers might be considered analogous to feminist or STEM-valuing mothers insofar as they are not highly gender traditional in gendered behavior or attitudes. Thus, they were also expected to use more supportive scaffolding (i.e., reading). Again I conducted a 3(scaffolding strategy: demonstrating vs. reading vs. correcting mistakes) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(work traditionality: high vs. low) mixed model ANOVA in which scaffolding strategy was a within-subjects variable and condition, gender, and work traditionality were between subjects variables. As with math importance and feminism, mothers were split into more and less traditional groups using a median split of their composite work scores (median = 1.17). Mauchley's test of sphericity indicated that the assumption of sphericity had been violated, $\chi^2(2) = 26.06$, $p < .001$, $\varepsilon = .72$, thus the Greenhouse-Geisser correction was applied. As before, there was a main effect of scaffolding type, $F(1.44,76.03) = 56.23$, $p < .001$, $\eta^2_p = .52$, and a scaffolding by child gender interaction, $F(1.44,76.03) = 10.57$, $p < .001$, $\eta^2_p = .17$. In this model, reduced from the 5 x 2 x 2 model used to examine strategy use among all mothers, a significant three-way interaction among scaffolding type, gender, and condition also emerged, $F(1.44,76.03) = 4.94$, $p = .018$, $\eta^2_p = .09$. Post-hoc pairwise comparisons revealed that among dyads playing with *BobbyBlox*, mothers were especially gendered in their strategy use, demonstrating building to sons more than to daughters ($M = .30$ vs. .17, $SD = .20$ and .12, Sidak-adjusted $p = .001$) and reading to daughters more than to sons ($M = .42$ vs. .27, $SD = .14$ and .19, $p = .001$). This interaction is depicted in Figure 10. There was a three-way interaction among scaffolding strategy, child gender, and mothers' work traditionality, $F(1.44,76.03) = 4.94$, $p = .023$, $\eta^2_p = .08$. Post-hoc comparisons revealed that, in partial support of the hypothesis, highly gender traditional mothers were especially gendered in their strategy use. Again, these mothers demonstrated building more to sons than to daughters ($M = .32$ vs. .16, $SD = .23$ and .11, Sidak-adjusted $p = .001$) and read to daughters more than to sons ($M = .43$ vs. .27, $SD = .12$ and .19, $p < .001$). This interaction is depicted in Figure 11. There were no other main effects or interactions.

**Children's Free Play**

Children were expected to be sex-typed in their mode of play, with girls engaging in more book use and symbolic play than boys, and boys engaging in more construction play than girls. Girls were expected to play more with *GoldieBlox* and boys were expected to play more with *BobbyBlox*. To examine children's play behaviors during free play, I conducted a 4(play mode: book building vs. other building vs. pretend play vs. not playing) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) mixed model ANOVA in which play mode was a within-subjects variable and condition and child gender were between-subjects variables. Mauchley's test of sphericity indicated that the assumption of sphericity had been violated, $\chi^2(5) = 74.85$, $p < .001$, $\varepsilon = .60$, thus the Greenhouse-Geisser correction was applied. The analysis showed a main effect of play mode, $F(1.78,98.12) = 22.47$, $p < .001$, $\eta^2_p = .29$. Post-hoc pairwise comparisons showed that building related and unrelated to the book were the most common play activities (Sidak-adjusted $p = .988$ between the two types of building). Not playing at all was the next most common activity (Sidak-adjusted *p*s < .001 between not playing and either type of building). Pretend play was the least common activity (Sidak-adjusted $p = .045$ between pretend play and not playing). Figure 12 depicts children's proportional modes of play, split by gender. There was a play mode by gender interaction, in partial support of the hypothesis that play would be sex-typed, $F(1.78,98.12) = 3.39$, $p = .043$, $\eta^2_p = .06$. Post-hoc comparisons showed that girls engaged in more book-related building than did boys ($M = .48$ vs. .29, $SD = .37$ and .34, Sidak-adjusted $p = .052$). Boys engaged in more construction unrelated to the book than did girls ($M = .52$ vs. .34, $SD = .32$ and .33, $p = .035$). Finally, there was a three-way interaction among play mode, gender, and condition, $F(1.78,98.12) = 3.39$, $p = .048$, , $\eta^2_p = .05$. Post-hoc comparisons showed that with *BobbyBlox* especially, girls used the book to guide building more so than boys did ($M = .52$ vs. .26, $SD = .34$ and .34, Sidak-adjusted $p = .050$). Boys built without the book more so than girls did ($M = .62$ vs. .24, $SD = .34$ and .22, $p = .001$). See Figure 13 for a graphic depiction of the two- and three-way interactions.

**GSF and free play.** Children who were especially gender salient were expected to be more gendered in their mode of play. To examine the role of personal gender salience in shaping children's mode of play, I conducted a 2(play mode: book building vs. other building) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(GSF: high vs. low) mixed model ANOVA in which play mode was a within-subjects variable and condition, child gender, and GSF were between-subjects variables. Children were divided into high- (H-GSF) and low-salience (L-GSF) groups using a median split of their composite GSF scores (median = -0.14). The children included in this analysis were those who completed the GSF measures in the post-test survey ($n$ = 35). As with the larger model, there was a significant gender by play mode interaction, with girls engaging in more book-related building than boys and boys engaging in more unrelated building than girls, $F(1,27) = 4.21$, $p = .050$, $\eta^2_p = .14$. This smaller model, reduced from the 5 x 2 x 2 model examining all models of play, also revealed a significant interaction between gender and condition, $F(1,27) = 8.24$, $p = .008$, $\eta^2_p = .23$. Girls engaged in more of both types of building with *GoldieBlox* than *BobbyBlox* ($M$ = .46 vs. .39, $SD$ = 0.34 and 0.24). Boys engaged in more of both types of building more often with *BobbyBlox* than with Goldie ($M$ = .47 vs. .29, $SD$ = .32 and .36). This provides limited support for the hypothesis that girls would play more with *GoldieBlox* and boys would play more with *BobbyBlox*. See Figure 14 for a graphic depiction of this interaction. There was no main effect of GSF, nor any interactions.

**Favorite toys and free play.** To permit examining the effect of children's gendered cognition on play behavior for the full sample, I also examined mode of play in relation to children's reported favorite toys at home. A more feminine toy at home might be associated with more feminine play behaviors and interest, for example. I conducted a 2(play mode: book building vs. other building) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(favorite toy: feminine vs. masculine) mixed model ANOVA in which play mode was a within-subjects variable and condition, child gender, and favorite toy were between-subjects variables. Children were split into feminine average and masculine average groups based on a median split of their favorite toy's gender rating (median = 2.99). Low scores reflected favorite toys that received feminine ratings while high scores reflected favorite toys that received masculine ratings. Analysis revealed a significant interaction between condition and children's toys, $F(1,51) = 13.86$, $p < .001$, $\eta^2_p = .21$. Children with a feminine favorite toy played more with *GoldieBlox* than children with a masculine favorite toy ($M = .47$ vs. .39, $SD = .38$ and .34). Children with a masculine favorite toy played longer with *BobbyBlox* than children with a feminine favorite toy ($M = .45$ vs. .24, $SD = 0.36$ and 0.30). There was also a three-way interaction among gender, condition, and children's toys, $F(1,51) = 5.98$, $p = .018$, $\eta^2_p = .11$. Post-hoc comparisons revealed that girls played more with *GoldieBlox* than *BobbyBlox* when they had a feminine favorite toy ($M = .45$ vs. .35, $SD = .41$ and .40, Sidak-adjusted $p = .041$). Boys also played more with *GoldieBlox* than *BobbyBlox* when they had a feminine favorite toy ($M = .49$ vs. .13, $SD = .26$ and .31, $p = .005$). Boys played more with *BobbyBlox* than *GoldieBlox* when they had a masculine favorite toy ($M = .46$ vs. .35, $SD = .34$ and .30, $p = .009$). These two- and three-way interactions are depicted in Figure 15. These findings are partially supportive of the hypothesis that girls would play more with *GoldieBlox* and boys would play more with *BobbyBlox*, in interaction with personal gender salience. There were no additional main effects or interactions.

**Interest in *GoldieBlox* and *BobbyBlox***

Similar to expectations about sex-typed play, children were expected to have sex-typed preferences for the game. Girls were expected to prefer *GoldieBlox* whereas boys were expected to prefer interested in *BobbyBlox*. I conducted a 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) ANOVA in which condition and child gender were between-subjects variables and children's rating of the game they played (Goldie vs. Bobby) was the dependent measure. There was a main effect of child gender, $F(1,57) = 10.38$, $p = .002$, $\eta^2_p = .15$. Girls were more interested in both games ($M = 2.71$, $SD = 0.46$) than were boys ($M = 2.23$, $SD = 0.68$). No main effect of condition or interaction between gender and condition was found. Children's preference for the games is depicted in Figure 16.

To further probe children's interest in the games, I compared their interest in having *GoldieBlox* versus *BobbyBlox* at home versus at school. I conducted a 2(location: home vs. school) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) mixed model ANOVA in which location was a within-subjects variable and condition and child gender were between-subjects variables. There was a marginally significant three-way interaction among location, child gender, and condition, $F(1,56) = 3.30$, $p = .075$, $\eta^2_p = .06$. Post-hoc comparisons revealed that boys were less interested in having GoldieBlox at school than were girls ($M = 1.86$ vs. 2.73, $SD = 0.86$ and 0.46, Sidak-adjusted $p = .003$). Girls were also more interested in having BobbyBlox at home than were boys ($M = 2.44$ vs. 1.87, $SD = 0.73$ and 0.83, $p = .032$). There was also a main effect of gender, $F(1,56) = 11.05$, $p = .002$, $\eta^2_p = .17$, whereby girls again expressed more interest in playing with both toys than did boys. Both of these effects are depicted in Figure 17. There were no other main effects or interactions.

**GSF and Blox interest.** To examine the role of GSF in shaping preference for the games, I conducted a 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(GSF: high vs. low) ANOVA in which condition, child gender, and GSF level were between-subjects variables, and children's rating of the game they played (Goldie vs. Bobby) was the dependent measure. Analysis did not reveal a main effect of GSF nor any interactions.

I also conducted a 2(location: home vs. school) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(GSF: high vs. low) mixed model ANOVA in which location was a within-subjects variable and condition, child gender, and GSF were between-subjects variables. As with the larger model with the full sample, analysis again showed a three-way interaction among location, gender, and condition, $F(1,27) = 7.74$, $p = .010$, $\eta^2_p = .22$. There was not a main effect of GSF nor any additional interactions, however.

**Favorite toys and Blox interest.** As with children's mode of play, favorite toys were used an additional indicator of gendered cognition for the full sample. I conducted a 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(favorite toy: feminine vs. masculine) ANOVA in which condition, child gender, and favorite toy were between-subjects variables, and children's rating of the game they played (Goldie vs. Bobby) was the dependent measure. As before, there was a main effect of child gender such that girls preferred both toys, $F(1,53) = 15.53$, $p < .001$, $\eta^2_p = .23$. There was a marginal main effect of favorite toy, $F(1,53) = 3.59$, $p = .063$, $\eta^2_p = .06$. Children with a masculine favorite toy rated *Goldie/BobbyBlox* (collapsed) more highly than did children with a feminine favorite toy ($M = 2.50$ vs. 2.11, $SD = 0.60$ and 0.65). There was an interaction between gender and favorite toy, $F(1,53) = 3.94$, $p = .050$, $\eta^2_p = .07$. Boys with a masculine favorite toy rated *Goldie/BobbyBlox* (collapsed) more highly than did boys with a feminine favorite toy ($M = 2.30$ vs. 1.50, $SD = 0.60$ and 0.50).

I also conducted a 2(location: home vs. school) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(favorite toy: feminine vs. masculine) mixed model ANOVA in which location was a within-subjects variable and condition, child gender, and favorite toy were between-subjects variables. Analysis revealed neither a main effect of favorite toy nor any interactions.

**Mothers interest in *GoldieBlox* and *BobbyBlox*.** Mothers' own perceptions about the game are relevant to understanding how they might guide play. To examine how much mothers enjoyed play with either game, I conducted a 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) ANOVA in which condition and child gender were between-subjects variables and mothers' rating of their enjoyment of the game they played (Goldie vs. Bobby) was the dependent measure. Analysis did not reveal a main effect or any interactions; mothers rated their enjoyment equally high in all cases ($M$ = 3.30, $SD$ = 0.56, possible range from 1-4). I also conducted a 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) ANOVA in which condition and child gender were between-subjects variables and mothers' rating of the game's educational value was the dependent measure. Again, analysis failed to reveal either the main effect or any interactions. Mothers rated the game equally educational in all cases ($M$ = 3.49, $SD$ = 0.54, possible range from 1-4).

Because mothers' perceptions could be influenced by their own values or gender attitudes, I included math importance (high vs. low) and work traditionality (high vs. low) in a series of additional analyses. First, I conducted a 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(math importance: high vs. low) ANOVA in which condition, child gender, and math importance were between-subjects variables and mothers' game enjoyment was the dependent measure. Analysis revealed a significant three-way interaction among gender, condition, and math importance, $F(1,36) = 5.40$, $p = .026$, $\eta^2_p = .13$. Post-hoc pairwise comparisons show that for mother-son dyads playing with *GoldieBlox*, mothers who valued math more enjoyed the game less than mothers who valued math less ($M = 3.12$ vs. 4.00, $SD = 0.64$, 0.00, Sidak-adjusted $p = .036$). See Figure 18 for a graphical depiction of this interaction.

Next, I conducted a 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(math importance: high vs. low) ANOVA in which condition, child gender, and math importance were between-subjects variables and mothers' rating of game educational value was the dependent measure. Analysis revealed a marginal three-way interaction among gender, condition, and math importance, $F(1,36) = 3.74$, $p = 0.62$, $\eta^2_p = .09$. For dyads playing with *GoldieBlox*, low math importance mothers rated it more educational for sons than for daughters ($M = 4.00$ vs. 3.00, $SD = 0.00$, 0.00, Sidak-adjusted $p = .042$). This interaction is depicted in Figure 19.

To examine the influence of gendered behavior and attitudes for the full sample of mothers, I repeated the previous two analyses using work traditionality as the individual differences variable of interest. First, I conducted a 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(work traditionality: high vs. low) ANOVA with condition, child gender, and work traditionality as between-subjects variables and mothers' game enjoyment as the dependent measure. There was a marginally significant interaction between condition and mother's work traditionality, $F(1,53) = 3.07$, $p = .080$, $\eta^2_p = .06$. Less traditional mothers enjoyed *BobbyBlox* more than *GoldieBlox* ($M = 3.61$ vs. 3.16, $SD = 0.51$ and 0.55). This interaction is depicted in Figure 20.

I next conducted a 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(work traditionality: high vs. low) ANOVA with condition, child gender, and work traditionality as between-subjects variables and mothers' rating of game educational value as the dependent measure. There were no significant main effects or interactions.

**Mechanical Learning**

Because children were expected to be sex-typed in their play and interest and because mothers were expected to be differentially supportive in their scaffolding of play, children were expected to have learning outcomes that varied in interaction with gender and condition. Girls were expected to learn more playing with *GoldieBlox* than with *BobbyBlox* whereas boys were expected to learn more from *BobbyBlox* than from *GoldieBlox*. Mechanical learning was assessed using a test of belt-drive learning. Application or transfer of that learning to an easier and a harder task drawing on similar principles was a secondary measure. To examine children's mechanical learning, I first conducted a 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) ANCOVA controlling for child age, with mechanical learning as the dependent variable. As expected based on preliminary analysis, there was a significant effect of the covariate, age, $F(1,56) = 7.81$, $p = .007$, $\eta^2_p = .12$. Older children outperformed younger children on the mechanical test. There was also a significant interaction between gender and condition, $F(1,53) = 5.27$, $p = .025$, $\eta^2_p = .09$. Contrary to expectations, girls learned more playing with *BobbyBlox* than with *GoldieBlox* ($M = 2.76$ vs. 1.92, $SD = 1.45$ and 0.80) and boys learned more playing with *GoldieBlox* than with *BobbyBlox* ($M = 1.85$ vs. 2.49, $SD = 1.67$ and 1.25). This interaction is depicted in Figure 21.

To test the transfer of mechanical learning to another context, I conducted a series of additional 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) ANCOVAs controlling for child age, first with time to complete the easy transfer task as the dependent variable and then score on the difficult task as the dependent variable. Because of the skew in scores, a log transformation was applied to help normalize the easy transfer task data. For the easy task, analysis revealed a significant effect of age, with older children outperforming younger children, $F(1,56) = 19.62$, $p < .001$, $\eta^2_p = .26$. There was a significant main effect of child gender, $F(1,56) = 4.15$, $p = .046$, $\eta^2_p = .07$. On average, boys completed the easy task faster than did girls ($M$ = 184.52 vs. 226.53 seconds, $SD$ = 92.21 and 108.85). For the difficult task, analysis again revealed a significant effect of age, with older children performing better than younger children, $F(1,54) = 14.69$, $p < .001$, $\eta^2_p = .21$. There were no other main effects or interactions. Performance on these tasks is depicted graphically in Figures 22 (easy task) and 23 (difficult task).

**GSF and mechanical learning.** Gender salience was expected to affect mechanical learning because GSF is considered to be a filter for attention (Liben & Bigler, 2002). If children are paying more attention to certain aspects of play, they may learn better or worse depending on the object of their attention. To examine the role of GSF in mechanical learning, I conducted a 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(GSF: high vs. low) ANCOVA controlling for child age, with mechanical learning as the dependent variable. Analysis again revealed a significant effect of age, $F(1,27) = 8.92$, $p = .006$, $\eta^2_p = .25$. There was a significant interaction between gender and GSF level, $F(1,27) = 8.74$, $p = .006$, $\eta^2_p = .07$. H-GSF girls outperformed L-GSF girls regardless of condition ($M$ = 2.81 vs. 1.74, $SD$ = 1.24 and 0.84). L-GSF boys outperformed H-GSF boys regardless of condition (M = 2.50 vs. 1.56, $SD$ = 1.35 and 0.98). Finally, there was a significant three-way interaction among gender, condition, and GSF level, $F(1,27) = 7.61$, $p = .010$, $\eta^2_p = .22$. Post-hoc comparisons showed that H-GSF girls learned significantly more from *BobbyBlox* than from *GoldieBlox* ($M$ = 3.80 vs. 1.81, $SD$ = 1.34 and 0.82, Sidak-adjusted $p = .002$). L-GSF boys also learned more from *BobbyBlox* than from *GoldieBlox*

($M$ = 3.16 vs. 1.85, $SD$ = 1.30 and 1.22, $p$ = .042). This interaction is depicted in Figure 24. There were no additional main effects or interactions.

**Favorite toys and mechanical learning.** As with earlier analysis, children's favorite toys were used as an additional indicator of gendered cognition to permit testing for the full sample. To examine the role of gendered toys in mechanical learning, I conducted a 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(favorite toy: feminine vs. masculine) ANCOVA controlling for child age, with mechanical learning as the dependent variable. As before, there was a main effect of child age, $F(1,52) = 6.19$, $p = .016$, $\eta^2_p = .11$ and an interaction between gender and condition, $F(1,52) = 7.73$, $p = .029$, $\eta^2_p = .09$. Supporting and partially replicating the pattern found with GSF, there was a three-way interaction among gender, condition, and favorite toys, $F(1,52) = 3.91$, $p = .053$, $\eta^2_p = .07$. Girls with a feminine favorite toy learned better from *BobbyBlox* than from *GoldieBlox* ($M$ = 3.15 vs. 1.83 , $SD$ = 1.14 and 0.88, Sidak-adjusted $p$ = .020). This interaction is depicted in Figure 25. There were no additional main effects or interactions.

**Parent work traditionality and mechanical learning.** Because parents' own gendered attitudes and behaviors may influence children's gendered attitudes and attention, they might be a distal influence on mechanical learning. To test this, both mothers' and fathers' work traditionality scores were used in the same ANCOVA set up from the previous analyses. I first conducted a 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(mother's work traditionality: high vs. low) ANCOVA controlling for child age, with mechanical learning as the dependent variable. As before, there was a main effect of child age, $F(1,52) = 4.54$, $p = .038$, $\eta^2_p = .08$ and an interaction between gender and condition, $F(1,52) = 12.31$, $p = .001$, $\eta^2_p = .19$. There was a main effect of mothers' work traditionality, $F(1,52) = 7.43$, $p = .024$, $\eta^2_p = .09$.

Mothers with less traditional work had children who performed better on the mechanical post-test than did mothers with more traditional work (*M* = 2.53 vs. 1.76, *SD* = 1.43 and 1.26). There was an interaction between mothers' work traditionality and child gender, $F(1,52) = 4.80$, $p = .033$, $\eta^2_p = .08$. Boys especially benefited from mothers with less gender-traditional jobs as compared to boys with traditionally-employed mothers (*M* = 2.64 vs. 1.15, *SD* = 1.50 and 1.29). Figure 26 depicts this interaction graphically. Finally, there was a marginally significant interaction between mothers' work traditionality and condition, $F(1,52) = 3.35$, $p = .073$, $\eta^2_p = .06$. Less traditional mothers playing with Bobby had children who did especially well on the mechanical post-test as compared to more traditional mothers and children playing with Bobby (*M* = 2.77 vs. 1.38, *SD* = 1.59 and 1.65). This interaction is depicted graphically in Figure 27.

Although fathers did not participate in the laboratory session, they likely still influence children's gender beliefs or mechanical aptitude. As with mothers, I conducted a 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(father's work traditionality: high vs. low) ANCOVA controlling for child age, with mechanical learning as the dependent variable. Fathers' work traditionality composite scores were split at the median into high and low traditionality groups (median = 0.17). As with the previous models, there was a main effect of child age, $F(1,52) = 12.98$, $p = .001$, $\eta^2_p = .20$, and a gender by condition interaction, $F(1,52) = 4.86$, $p = .032$, $\eta^2_p = .09$. There was a marginal main effect of fathers' work traditionality, $F(1,52) = 3.90$, $p = .054$, $\eta^2_p = .07$. Fathers with more traditional work had children who performed better on the mechanical post-test than did fathers with less traditionally masculine work (*M* = 2.58 vs. 1.93, *SD* = 1.43 and 1.26). There was also a three-way interaction among gender, condition, and father's work traditionality, $F(1,52) = 4.06$, $p = .049$, $\eta^2_p = .07$. Girls whose fathers had gender-traditional jobs performed especially well with *BobbyBlox* as compared to *GoldieBlox* (*M* = 3.28

vs. 1.97, *SD* = 1.25 and 0.89, Sidak-adjusted $p = .043$). Boys whose fathers had gender-traditional jobs performed especially well with *GoldieBlox* as compared to *BobbyBlox* ($M = 3.26$ vs. 1.83, *SD* = 1.55 and 1.57, $p = .028$). Figure 28 depicts this interaction graphically. There were no additional main effects or interactions.

**Interest in Other STEM Toys**

The final outcome of interest was interest in other STEM toys. Apart from learning mechanical skill, children might become more interested in other STEM toys following play with a particularly fun STEM toy. I expected that children would be sex-typed in their interest in other STEM toys but that girls who played with *GoldieBlox* would be more interested in other STEM toys overall than girls who played with *BobbyBlox*. To test the effect of game play on children's interest, I conducted a 3(STEM toy type: feminine vs. neutral vs. masculine) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) ANCOVA controlling for child age, with STEM toy type as a within-subjects variable and child gender and condition as between-subjects variables. Mauchley's test of sphericity indicated that the assumption of sphericity had been violated, $\chi^2(2) = 10.12$, $p = .006$, $\varepsilon = .86$, thus the Huynd-Feldt correction was applied. Analysis revealed a significant STEM toy type by gender interaction, $F(1.86, 105.89) = 47.78$, $p < .001$, $\eta^2_p = .46$. Post-hoc pairwise comparisons reveal that, as expected, children's interests were sex-typed. Girls were most interested in feminine toys ($M = 2.74$, $SD = 0.29$, Sidak-adjusted *p*s $< .001$), then neutral toys ($M = 2.28$, $SD = 0.53$), and least interested in masculine toys ($M = 2.04$, $SD = 0.58$, $p = .007$ between neutral and masculine). Boys were most interested in masculine toys and neutral toys ($M = 2.52$ and 2.41, *SD* = 0.39 and 0.47, $p = .418$ between masculine and neutral) and least interested in feminine toys ($M = 1.91$, $SD = 0.53$, *p*s $< .001$). This interaction is depicted in Figure 29. There was also a significant between-subjects main effect of the covariate, age, $F(1,56) = 7.32$, $p < .001$, $\eta^2_p = .12$, where older children were more interested in the STEM toys.

**GSF and other STEM toys.** GSF was potentially expected to affect interest in other STEM toys in similar ways as it was expected to affect interest in *GoldieBlox* or *BobbyBlox*, with H-GSF children being especially sex-typed in their preferences. To test the role of GSF in shaping interest in other STEM toys, I conducted a 3(STEM toy type: feminine vs. neutral vs. masculine) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(GSF: high vs. low) ANCOVA controlling for child age, with STEM toy type as a within-subjects variable and child gender, condition, and GSF as between-subjects variables. Analysis did not reveal a main effect of GSF nor any additional interactions.

**Favorite toys and STEM toys.** As with previous analyses, favorite toys were used as an indicator of gendered cognition for the full sample. To test the role of children's favorite toys in shaping interest in future play with other STEM toys, I conducted a 3(STEM toy type: feminine vs. neutral vs. masculine) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) x 2(favorite toy: feminine vs. masculine) ANCOVA controlling for child age, with STEM toy type as a within-subjects variable and child gender, condition, and favorite toy as between-subjects variables. Mauchley's test of sphericity indicated that the assumption of sphericity had been violated, $\chi^2(2) = 8.74$, $p =.013$, $\varepsilon = .86$, thus the Huynd-Feldt correction was applied. As in the original model, analysis revealed a significant interaction between STEM toy type and gender, $F(2.00,104.00) = 15.82$, $p < .001$, $\eta^2_p = .23$, and a between-subjects main effect of child age, $F(1,52) = 5.27$, $p = .026$, $\eta^2_p = .09$. There was a significant main effect of favorite toy, $F(1,52) = 12.62$, $p = .001$, $\eta^2_p = .20$. Children with a masculine favorite toy were more interested in STEM toys overall ($M = 2.44$, $SD = 0.48$) than were children with a feminine favorite toy ($M = 2.03$, $SD$

= 0.51). This model also revealed a significant main effect of gender, $F(1,52) = 8.27$, $p = .006$, $\eta^2_p = .14$. Girls were more interested in all toys ($M = 2.40$, $SD = 0.47$) than were boys ($M = 2.07$, $SD = 0.46$). There were no additional main effects or interactions.

**Mothers' Interest in Purchasing STEM Toys.** Because parents often determine children's access to particular toys, mothers were also queried about their interest in purchasing the three types of STEM toys (feminine, neutral, masculine). To test mothers' interest in buying different STEM toys for their children, I conducted a 3(STEM toy type: feminine vs. neutral vs. masculine) x 2(condition: Goldie vs. Bobby) x 2(child gender: girl vs. boy) mixed model ANOVA with STEM toy type as a within subjects variable and condition and child gender as between subjects variables. Analysis showed a main effect of STEM toy type, $F(2,114) = 24.76$, $p < .001$, $\eta^2_p = .30$. Pairwise comparisons showed that mothers were most interested in purchasing neutral STEM toys ($M = 3.09$, $SD = 0.46$, Sidak-adjusted *p*s < .001), then masculine STEM toys ($M = 2.81$, $SD = 0.60$). Mothers were least interested in purchasing feminine-marketed STEM toys ($M = 2.58$, $SD = 0.52$, $p = .011$ between feminine and masculine). There was also a STEM toy type by child gender interaction, $F(2,114) = 22.89$, $p < .001$, $\eta^2_p = .29$. For daughters, mothers were most interested in purchasing neutral toys ($M = 3.08$, $SD = 0.49$), followed by feminine toys ($M = 2.81$, $SD = 0.50$, Sidak-adjusted $p = .040$ compared to neutral), and were least interested in purchasing masculine toys ($M = 2.54$, $SD = 0.62$, $p = .046$ compared to feminine, < .001 compared to neutral). For sons, mothers were most interested in purchasing neutral or masculine toys ($M = 3.11$ and 3.08, $SD = 0.43$ and 0.44, Sidak-adjusted $p = .987$ between neutral and masculine) and least interested in purchasing feminine toys ($M = 2.36$, $SD = 0.44$, *p*s < .001). This interaction is depicted in Figure 30.

# Chapter 4

# DISCUSSION

The present study was conducted to answer three related research questions: (1) Do mothers guide play with a STEM toy differently with daughters versus with sons? (2) Do girls and boys approach and use STEM toys differently, irrespective of whether those toys are explicitly marketed or designed to appeal to girls versus boys? (3) How is play affected if the toy is marketed explicitly to girls versus to boys? I begin the discussion of results with a brief summary of the findings in broad stroke, and then return to each of these research questions in turn. I then address the role of (a) children's individual differences and (b) parents' individual differences in shaping play and mechanical learning outcomes. I conclude by discussing implications for developmental theory and STEM intervention, and make suggestions for future research in this area.

## Summary of Major Findings

From the outset, marketing the toy as "for girls" versus "for boys" impacted play. Mothers engaged in more building during toy familiarization with *BobbyBlox* than with *GoldieBlox*. Dyads engaged in a variety of play behaviors with the toy, including building along with the book, building without the book, reading the book, among others. Mothers differentially scaffolded play depending on the child's gender, with mothers using reading as a tool more often with daughters than with sons, and demonstrating building more often with sons than with daughters. Importantly, reading the book was the only scaffolding strategy positively associated with children's mechanical learning. In contrast, correcting children's mistakes and demonstrating building were both predictive of completing the game during the dyadic interaction. Game completion and mechanical learning may be considered two distinct goals.

During free play, children's mode of play interacted with gender. Girls engaged in more book-related building whereas boys engaged in more building unrelated to the book. This was especially true in the *BobbyBlox* condition, where girls relied on the book to guide building and boys did not often use the book. Girls learned the belt-drive concept better from *BobbyBlox* and boys learned better from *GoldieBlox*. Mechanical learning did not translate to a transfer task. Interest in play with future STEM toys was sex-typed, as expected. Mothers' interest in purchasing STEM toys was largely sex-typed, though they preferred neutral toys for both sons and daughters.

**Do Mothers Guide Play with a STEM Toy Differently with Daughters Versus Sons?**

Mothers indeed guided play differently with daughters versus sons. Mothers' use of scaffolding strategies interacted with child gender, accounting for 17% of the variance in strategy use in the 3 x 2 x 2 x 2 mixed model ANOVA examining scaffolding by condition, gender, and mothers' work traditionality. Mothers read the book more often with daughters than with sons, and demonstrated building more often with sons than with daughters. Reading the book was the most common scaffolding strategy, with the largest mean relative frequency, and the only strategy associated with mechanical learning. Demonstrating building was a somewhat less common strategy relative to reading, and was associated with game completion but not learning (along with correcting mistakes). Correcting mistakes was also associated with game completion but not learning and mothers used this strategy equally with sons and daughters.

It may be that mothers had different goals for daughters versus for sons. For example, it may be the case that mothers were attempting to keep sons on-task and thus employed strategies geared toward game completion. In contrast, if daughters were more focused, mothers may have employed other strategies geared toward learning. Future video coding should try to disentangle the role of child behavior in eliciting particular maternal strategies. Interestingly, there was no gender difference in how often mothers corrected children's mistakes, though this strategy was positively associated with game completion. Also of note is that there was no gender difference overall in mechanical learning or in rate of game completion. Still, dyads were not given specific instructions with the game. It is possible that mothers had different goals with daughters versus with sons (e.g., process vs. outcome), whether or not those are tapped by the game completion or mechanical learning outcome measures.

Certain types of mothers appeared to be especially affected by child gender. Mothers who were involved in highly gender-traditional occupations were especially gender-typed in their scaffolding, reading with daughters but not sons and demonstrating building with sons but not daughters. This three-way interaction among mothers' work traditionality, child gender, and scaffolding strategy accounted for an additional 8% of the variance in strategy use. It appears that mothers' work traditionality might have translated to behavior in the laboratory session. Work traditionality was used here as an indicator both of gendered cognition (e.g., gendered interest) and gendered behavior. These distinct processes (cognition and behavior) may have influenced scaffolding though a number of mechanisms. More traditional mothers might have been more practiced in sex-typed behavior, although this does not account for their demonstrating building to sons, as building is stereotypically masculine. Traditional mothers might have more traditional gender role attitudes and thus intentionally or unconsciously apply different strategies with sons versus with daughters. There were other mother individual difference variables that impacted scaffolding (i.e., feminism and math importance), and I return to those in the section pertaining to the effect of marketing.

## Do Girls and Boys Approach and Use STEM Toys Differently, Irrespective of Marketing?

Girls and boys differed on average in their mode of play with the toys and in the amount of time spent playing in general. The most common activities were building along with the book or building unrelated to the book, which differed by gender. The gender by mode of play interaction accounted for 6% of the variance in the 4 x 2 x 2 mixed model ANOVA comparing play mode by gender and condition. Girls engaged in more book-related building than boys did. Boys engaged in more building unrelated to the book than girls did. Book-related building may reflect a more feminine approach to using the toy given that it incorporates reading the book. Or, it might reflect compliance or following directions. In contrast, building without the book may reflect a more masculine approach to using the toy given that it involves construction without book guidance. Boys, having more construction play experience than girls (e.g., Jirout & Newcombe, 2015), may simply be more comfortable building without the book's guidance. Yet, it is important to note that construction play is associated with spatial skill. The belt-drive, a *mechanical* concept, was likely novel for both boys and girls. Given that the book teaches the belt-drive, abandoning the book may ultimately not have been advantageous for mechanical learning. It is also important to note that while there were mean gender differences, there was significant overlap in mode of play as well.

## How is Play Affected if the Toy is Marketed Explicitly to Girls versus to Boys?

Marketing mattered in the present study in multiple, complex ways. It affected mothers' initial approach to the toy, some mothers' scaffolding, children's play, and ultimately, children's learning. Marketing of other STEM toys was also associated with sex-typed interests.

**Marketing affects maternal scaffolding.** From the outset, marketing mattered. Mothers prepare differently when given *BobbyBlox* than when given *GoldieBlox*, building more with the former than the latter. In dyadic play, mothers were affected by children's gender, employing different strategies with daughters versus sons. The smaller 3 x 2 x 2 x 2 mixed model ANOVA examining scaffolding use by child gender, condition, and mothers' work traditionality also showed that mothers playing with *BobbyBlox* were particularly differentiated in their scaffolding between boys and girls. It may be the case that because *BobbyBlox* is marketed to boys, mothers demonstrate building with boys but focus on reading (i.e., following the instructions) with girls.

Some mothers were differentially affected by the marketing of the toy. Mothers who endorsed feminist values were especially sex-typed in their scaffolding with *GoldieBlox*, reading more with daughters and demonstrating building more with sons. In contrast, mothers who did not endorse feminist values were most sex-typed in their scaffolding when the toy was *BobbyBlox* rather than *GoldieBlox*. The interaction of mothers' feminism, child gender, condition, and scaffolding strategy type accounted for 16% of the variance in strategy use. In a consumer marketing survey given to purchasers of the first production batch of *GoldieBlox and the Spinning Machine*, mothers identified feminism as one motivation for their purchase. It may be that some feminist mothers are attracted to the idea of a STEM toy explicitly marketed to girls and thus employ reading as a strategy with daughters because they are focused on the belt-drive instructions and oriented toward mechanical learning.

**Marketing affects children's play.** Children's free play was affected by game marketing. With *BobbyBlox*, girls engaged in building guided by the book much more so than boys, whereas boys engaged in building without the book more so than girls. It may be that because *BobbyBlox* is marketed to boys and because boys play more with construction toys than do girls, boys approach the toy with the assumption they know how to use it. Thus, boys do not use the book to guide play as much as girls do with either set, but especially not with *BobbyBlox*. In contrast, a girl might approach *BobbyBlox* as an unfamiliar toy and use the book as a critical set of directions to support and guide play.

Children with a feminine favorite toy built more (both types) with *GoldieBlox* than *BobbyBlox*, whereas children with a masculine favorite toy built more with *BobbyBlox* than with *GoldieBlox*. These findings support the role that previous experience plays in shaping approach to new experiences. Children with feminine favorite toys, those that consider such toys as being "for me," likely also consider the feminine *GoldieBlox* to be such a toy and approach and build with it more so than the masculine *BobbyBlox*. The same is true in reverse for children with masculine favorite toys: *BobbyBlox* is "for me" and thus approachable whereas *GoldieBlox* is not.

**Marketing affects some children's and mothers' toy interest.** Marketing did not appear to greatly affect either children's or mother's reported interest in the toy on average, with some specific exceptions. Overall, mothers rated both sets equally enjoyable and educational. However, mothers who highly valued math enjoyed *GoldieBlox* less than mothers who did not value math highly. Mothers with less gender traditional work enjoyed *GoldieBlox* less than *BobbyBlox*. Additionally, mothers who did not value math highly rated *GoldieBlox* as less educational for sons than for daughters. It may be that toys marketed to girls are perceived as less interesting or educational to certain parents and for certain children. In this case, *GoldieBlox* was less appealing to mothers who valued math (for sons and daughters), mothers who did not especially value math but were potentially shopping for sons, and mothers who themselves have less traditional gender role attitudes or behavior.

Girls enjoyed both *GoldieBlox* and *BobbyBlox* more than boys did. Gender accounted for 15% of the variance in interest. When asked to rate interest in having either toy at home versus at school, boys were somewhat less interested in having *GoldieBlox* at school than were girls, whereas girls were more interested in having *BobbyBlox* at home than were boys, accounting for an additional 6% of the variance in interest. It may be that boys are wary of playing with a feminine toy in the presence of their peers. Future research should examine the role of peers in shaping approach to, and play with a STEM toy given that play with peers is generally sex-typed in mode.

Interestingly, though girls liked both toys more than boys did, children with a masculine favorite toy were more interested in *Goldie/BobbyBlox* (collapsed) than were children with a feminine favorite toy. Girls' interest in both toys might reflect a preference for feminine aspects of the toy design (e.g., animals, book) but this does not align with the finding that children with masculine favorite toys were also more interested. In fact, girls' preference for both toys may reflect compliance or agreeableness on some level, in addition to enjoyment of particular design aspects of the toys.

**Marketing affects mechanical learning and STEM interest.** One of the strongest pieces of evidence that marketing meaningfully impacts STEM toy outcomes was the finding that girls learned more playing with *GoldieBlox* than *BobbyBlox* and that boys learned more from *BobbyBlox* than *GoldieBlox*. Gender accounted for almost 10% of the variance in mechanical learning (on top of age, which accounted for 12%). This interaction is quite striking considering that girls built more overall with *GoldieBlox* and boys built more overall with *BobbyBlox*. Evidence from dyadic and free play sheds some light on the differential learning outcomes. Because girls used the book to guide building and boys built without the book, especially with *BobbyBlox*, it may be the case that with *BobbyBlox*, boys learned less because they relied less on the book as a set of instructions and girls learned more because they indeed did rely on the book. With *GoldieBlox*, girls are doing relatively equal levels of book building and building without the book, suggesting that like boys with *BobbyBlox*, they may perceive they know how to use an own-gender toy and do not need directions. Because the belt-drive is a new concept to the boys and girls in this sample, and no child had used the toys before the lab session, all children would have been aided by the book in learning the mechanical concept. The gender by condition interaction did not transfer to performance on two transfer tasks, where the only effect was children's age.

Importantly, gender salience interacted with gender and condition to impact learning. H-GSF girls especially benefitted from *BobbyBlox* relative to *GoldieBlox*. L-GSF boys also learned more from *BobbyBlox* than from *GoldieBlox*. Gender salience accounted for an additional 22% of the variance in mechanical learning. This suggests that for H-GSF girls, the very girls who may be most attuned to the gender cues in a toy marketed by gender, marketing affects learning especially strongly. Similarly, children with a feminine favorite toy learned better from *BobbyBlox* than from *GoldieBlox*. Play with *GoldieBlox* may be less educational than play with *BobbyBlox* for the same girls that might be clamoring to buy it. There was no particular interaction for H-GSF boys; H-GSF boys perform with the same pattern overall as boys on average, doing better with *GoldieBlox* and worse with *BobbyBlox*. L-GSF boys looked very similar to H-GSF girls in terms of performance but rather than being an effect of GSF *per se*, it may have been the result of gendered interest given that these boys paid less attention to gender. The role of gender salience in STEM toy play is unpacked further in the upcoming subsection, Personal Gender Salience.

**Interest in future play with STEM toys.** Interest in other STEM toys was not affected by the marketing of the toy children played with (i.e., *GoldieBlox* vs. *BobbyBlox*) but it was affected by the marketing of the depicted STEM toy, accounting for 46% of the variance in interest. Girls were interested in playing with feminine toys more so than neutral toys, and neutral toys more so than masculine toys. In contrast, boys were interested in playing with masculine or neutral toys more so than feminine toys. This suggests that regardless of play condition, boys and girls gravitated back to sex-typed preferences for future toys. It is interesting to note that boys might view neutral and masculine STEM toys as belonging to the same category. This is partially informed by the finding that children with a masculine favorite toy liked *Goldie/BobbyBlox* (collapsed) better and rated all STEM toys (collapsed) more highly than children with a feminine favorite toy. It could potentially be the case that there is some overlap between neutral and masculine when it comes to perceptions of STEM toys because the category STEM is culturally stereotyped as masculine. When the toys were clearly advertised as being "for girls," girls rated them positively and boys rated them negatively.

Mothers, too, rated interest in purchasing future STEM toys along gender lines with one important exception: Mothers indicated most interest in purchasing neutral toys overall. Mothers were more interested in purchasing neutral than feminine toys for daughters, and equally interested in purchasing neutral and masculine toys for sons. The finding that mothers rated neutral and masculine STEM toys as equally desirable for sons is similar to boys' ratings of equal interest in both categories, again suggesting category overlap. Mothers' generally low rating of feminine-marketed toys, taken with the findings that certain mothers rated *GoldieBlox* as less enjoyable or educational, suggests this category of toy might be perceived as less desirable or educational overall.

**Personal Gender Salience**

Marketing played a key role in shaping mothers' and children's play, interest, and learning. Importantly, marketing operated in interaction with children's individual differences in gender salience and favorite toys, an indicator of gendered interest and behavior. These individual differences affected how children played, what they learned, and their interest in future play with STEM toys. As discussed, children with a feminine favorite toy built more with *GoldieBlox* than with *BobbyBlox*. Children with a masculine toy built more with *BobbyBlox* than with *GoldieBlox*. H-GSF girls, L-GSF boys, and children with a feminine favorite toy learned more playing with *BobbyBlox* than with *GoldieBlox*. Finally, children with a masculine favorite toy were more interested in STEM toys overall. These findings provide support for Liben and Bigler's dual pathways model (2002).

Previous research by Coyle and Liben (2013) found that H-GSF girls were more interested in feminine play activities following an intervention with Barbie as the primary character. Their work suggested that while hyper-feminized models like Barbie are not effective in increasing girls' interest in masculine jobs (including STEM), such an intervention approach appears actually to backfire for some girls. Similarly, in the present study, H-GSF girls learned little mechanical skill playing with *GoldieBlox*. In contrast, these girls performed very well after playing with *BobbyBlox*. I have already speculated that mothers might have been more effective teachers for girls and boys in the *BobbyBlox* condition and that boys and girls appear to have used the book to guide play less in their own-sex conditions. Those findings account for the gender by condition interaction but not for the additional interaction with gender salience or favorite toy. Strength of GSF, or literally, the degree to which children *filter* the world by gender (Liben & Bigler, 2002), moderated learning outcomes. Mode of play was not dramatically different for H-GSF girls as compared to girls as a group or to L-GSF girls. Children with a feminine favorite toy actually built more with *GoldieBlox* than with *BobbyBlox* and yet showed the same learning pattern as H-GSF girls, learning more from *BobbyBlox* than from *GolieBlox*.

Latent processes at the level of cognition may be responsible, rather than easily observable differences in mode of play. For H-GSF girls playing with *GoldieBlox*, there were numerous cues that this was a toy "for girls," their gender category. It may be that the feminine aspects of the toy consumed most of the H-GSF girls' attention, precluding much attention being paid to learning the mechanical concepts. Thus although their play appears similar to that of girls as a group in terms of the relative frequency of building using the book and building unrelated to the book, their learning outcomes differed.

Because GSF was observable only in the subset of the sample that completed the GSF measures, favorite toy was used as an additional indicator of gendered interest and behavior for the full sample. These findings were very similar to the GSF findings in many ways. Additionally, they speak to the role of experience in shaping interest. Children's sex-typed play during the preschool years is well-documented (Cherney et al., 2003; Jirout & Newcombe, 2015; e.g., Martin et al., 1995). Sex-typed play may also be self-perpetuating, with sex-typed play engendering more sex-typed play in the future (Martin et al., 2012). Thus, it is not surprising that children with a feminine favorite toy were more interested in the feminine-marketed *GoldieBlox* than the masculine-marketed *BobbyBlox* whereas children with a masculine favorite toy were more interested in *BobbyBlox* than in *GoldieBlox*. It is also perhaps not surprising that children with a feminine favorite toy built more with *GoldieBlox* than with *BobbyBlox* if that indicates comfort or fluency with feminine toys. However, it is somewhat more surprising that children with feminine favorite toys learned more from *BobbyBlox* than from *GoldieBlox*. It may be the case that similar cognitive processes are at work with these children as with the H-GSF girls – their interest in feminine toys might be distracting from learning the mechanical concepts when given a feminine toy, potentially familiar and appealing in many ways though novel in mechanics.

It is both interesting and distressing that children with a masculine favorite toy were more interested in STEM toys overall. As discussed previously, these same children were more interested in *Goldie/BobbyBlox* overall than were children with a feminine favorite toy. Taken together, those findings indicate that children did recognize *Goldie/BobbyBlox* as belonging to the category of STEM toys. They also suggest that STEM toys are differentially appealing as a category for children already attracted to or comfortable with masculine toys.

**The role of personal interest.** L-GSF boys performed similarly to L-GSF girls, learning more from *BobbyBlox* than from *GoldieBlox*. This is further support for the dual pathways model proposed by Liben and Bigler (2002). For low salience children, personal interest may matter more than gender *per se* (although the two are almost certainly related). These boys were may have been less distracted by the gender cues in *BobbyBlox* than H-GSF girls were with *GoldieBlox.* Yet, they were very interested in the toy given that it was explicitly marketed “for boys.” Again there were not observable differences in mode of play. At the level of cognitive processing, however, L-GSF boys were able to glean something from *BobbyBlox* that boys as a group did not. Thus for L-GSF boys but not for H-GSF boys, *BobbyBlox* was indeed beneficial for learning.

## Parents’ Gender Attitudes and Behavior

Just as children’s own gendered attitudes and behavior affected approach, play, and learning, so too did parents’ gendered attitudes and behavior. Individual differences in mothers’ gender beliefs and behaviors interacted with marketing to affect how mothers scaffolded play, how mothers responded to the toys, and what children learned from the toys. As discussed, feminist mothers used *GoldieBlox* in especially gender-typed ways whereas low feminist mothers used *BobbyBlox* that way. Although overall most mothers enjoyed *GoldieBlox* and *BobbyBlox* equally and rated them equally educational, mothers who valued math enjoyed *GoldieBlox* less than *BobbyBlox* when playing with a son. Similarly, less gender traditionally employed mothers enjoyed *BobbyBlox* more than *GoldieBlox*. Mothers who did not especially value math rated *GoldieBlox* as more educational for sons than for daughters. Importantly, both mothers' and fathers' work traditionality affected children's mechanical learning.

Children performed better when parents' work was more traditionally masculine (i.e., low gender traditional mother or high gender traditional father). Boys especially benefited from mothers with less gender traditional employment. This may be due in part to the finding that more traditional mothers were more sex-typed in their scaffolding and demonstrated building more with boys, a strategy associated with game completion but not mechanical learning. Children playing with *BobbyBlox* also benefited from a less gender traditionally employed mother, perhaps for similar reasons and perhaps also because less traditional mothers enjoyed *BobbyBlox* more. It could be the case that when mothers' enjoyed the game, they conveyed that enjoyment and engagement to their children in a way that was supportive of mechanical learning.

Finally, fathers' gender-traditional work was especially beneficial for girls with *BobbyBlox* and boys with *GoldieBlox*. Given that masculinity of both fathers' and mothers' work was beneficial, yet fathers were not present for the session, it is critical to consider what masculinity of work might mean in terms of hereditary mechanical skill, gender role attitudes in the home, and gendered behavior at home. On average, the families in the present study reported having adopted the *breadwinner-caregiver* model of work within and outside the home, with one parent responsible mostly responsible for childcare and the other parent largely responsible for income earning (see Fulcher & Coyle, 2011). The families in the present study reported holding gender traditional jobs, and splitting household labor and childcare on gender-traditional lines. Thus, masculinity of fathers' job reflects traditional employment and likely, traditional breadwinner-caregiver structure. In contrast, masculinity of mothers' job reflects nontraditionality. Yet, both resulted in similarly positive mechanical learning outcomes for those children.

It may be the case that fathers or mothers with more traditionally masculine jobs simply have more mechanical skill which they have in turn imparted on their children, through genes, direct tuition, or another mechanism. It could also be the case that fathers' work traditionality was beneficial for entirely different reasons than mothers' work traditionality was beneficial. Certainly, mothers were present for the laboratory session and thus active participants in children's learning whereas fathers' influence on learning was relatively distal, which was one limitation of the current study. Disentangling the mechanisms behind mothers' versus fathers' influence on children's mechanical learning should be tested explored further in future research that includes fathers in testing. Directly examining the proximal and distal role of each parent would better elucidate why the present study shows similar effects for aspects of mothers' and fathers' gender attitudes and behavior on learning.

**Implications for Developmental Theory and STEM Intervention**

The present study has important implications for developmental theory and for translation to STEM intervention. First, the present study tested the role of Liben and Bigler's (2002) GSF as a moderator of children's learning from a toy marketed to girls versus to boys. These results support previous findings by Coyle and Liben (2013) that GSF is a critical moderator of children's experience in gender-linked encounters. The present study builds on their work by demonstrating that GSF is important not only in shaping interest but also in affecting learning. Together these findings underscore the importance of assessing GSF as a routine part of intervention assessment. Findings also suggested that the second of the dual pathways, personal interest (Liben & Bigler, 2002), may shape outcomes independently of, or separately from, GSF. Future research should continue to assess GSF and personal interest to better understand the way the two pathways operate independently and in tandem.

Parents' individual differences were also important for children's learning outcomes although the present study was limited by only including mothers in the laboratory session. Further research with both partners is needed to understand the mechanism behind each parents' influence. It would be interesting in future research to examine the dyadic interaction between peers, GSF, and STEM interest or learning as well. The present study focused on parents for theoretical and practical reasons but peers are an increasingly critical part of children's social network as they age (Degner & Dalege, 2013; Harris, 1995, 2009; Rubin et al., 2006).

To date, STEM intervention has been approached in myriad ways, in a range of settings and with nearly all age groups (Liben & Coyle, 2014). Little is understood about what works well, what does not work, and what might even be harmful. Similar to recent work by Jirout and Newcombe (2015), the current study suggests some toys can indeed teach foundational skills relevant for STEM. However, the marketing of such toys was deeply impactful for learning in the current study. Taking together, the findings from this study and those from Coyle and Liben's earlier work (2013), demonstrate that using feminized toys to attract girls to STEM may not be successful or beneficial. Yet, marketing STEM as feminine or "for girls" happens in a variety of programs as wide-ranging as EYH and Science cheerleaders, as discussed in the Introduction (Cavalier, 2014; Expanding Your Horizons Network, 2013). Moreover and perhaps more practically relevant to the present study, *GoldieBlox* is an actual toy currently available for purchase, whereas *BobbyBlox* is not. In the United States, toys are marketed even more by gender now than they were historically (Auster & Mansbach, 2012; Sweet, 2014). Children's ratings of interest in future STEM toys suggest that regardless of what they might be given to play with, after play they still gravitate toward gender-marketed STEM toys, like *GoldieBlox* for girls and *BobbyBlox* for boys.

One potential lesson for intervention to be taken from the present study is not to market STEM as being "for girls." Marketing for either gender benefitted learning for one group of children while hindering learning for the other. Those children especially attuned to gender (H-GSF girls) were also hindered by play with feminine-marketed *GoldieBlox* as compared to *BobbyBlox*. In fact, neutral approaches to intervention, accessible to boys and girls alike, might be more successful. If there is less information to filter by gender and if the neutral environment is category-appropriate (e.g., "for girls and for boys" or "for everyone"), H-GSF girls might not be so caught up by attractive feminized cues. As discussed in the Introduction, work by Cheryan and colleagues (Cheryan et al., 2009) suggests that college women were more interested in computer science when surveyed in a classroom with neutral decoration than a classroom with stereotypically masculine decoration. Future research should look at children's use of, and learning from, a more neutral STEM toy. It is promising to note that mothers indicated most interest in purchasing neutral STEM toys overall, even though their children indicated sex-typed preferences. However, mothers' second choice in both cases was the gender-typed category (in the case of sons, neutral and masculine ratings were equal). It is an empirical question whether mothers would choose the neutral toys if their child requested the sex-typed toys. Moreover, toy use might vary depending on whether the play context is didactic, as with a parent or teacher, versus child-directed, as with free play or peer play.

**Conclusion**

In summary, the present study demonstrated the powerful effect of gender-based marketing on both mothers' and children's play with, interest in, and learning from a STEM toy. Maternal scaffolding and children's play were also shaped by children's gender. Play and learning outcomes interacted with both children's and parents' gender attitudes and behaviors. These findings call attention to the importance of assessing individual differences to understand and interpret intervention outcomes. From a STEM intervention perspective, the focus taken by some to market to girls exclusively or on the basis of attractive femininity specifically may be misguided. In the most gender-segregated toy market in U.S. history, neutral play options are scarce – yet such options appear to be of critical importance for children's earliest learning.

**REFERENCES**


Auster, C. J., & Mansbach, C. S. (2012). The gender marketing of toys: An analysis of color and type of toy on the Disney Store website. *Sex Roles, 67*(7-8), 375-388. doi: 10.1007/s11199-012-0177-8

Becker, J. C., & Wagner, U. (2009). Doing gender differently - The interplay of strength of gender identification and content of gender identity in predicting women's endorsement of sexist beliefs. *European Journal of Social Psychology, 39*(4), 487-508. doi: 10.1002/ejsp.551

Bem, S. L. (1983). Gender schema theory and its implications for child development: Raising gender-aschematic children in a gender-schematic society. *Signs, 8*(4), 598-616.

Blakemore, J. E. O. (2003). Children's beliefs about violating gender norms: Boys shouldn't look like girls, and girls shouldn't act like boys. *Sex Roles, 48*(9-10), 411-419. doi: 10.1023/A:1023574427720

Blakemore, J. E. O., & Centers, R. E. (2005). Characteristics of boys' and girls' toys. *Sex Roles, 53*(9-10), 619-633. doi: 10.1007/s11199-005-7729-0

Blakemore, J. E. O., & Hill, C. A. (2008). The child gender socialization scale: A measure to compare traditional and feminist parents. *Sex Roles, 58*(3-4), 192-207. doi: 10.1007/s11199-007-9333-y

Buijzen, M., Van Reijmersdal, E. A., & Owen, L. H. (2010). Introducing the PCMC model: An investigative framework for young people's processing of commercialized media content. *Communication Theory, 20*(4), 427-450. doi: 10.1111/j.1468-2885.2010.01370.x

Carbonaro, M., Szafron, D., Cutumisu, M., & Schaeffer, J. (2010). Computer-game construction: A gender-neutral attractor to Computing Science. *Computers & Education, 55*(3), 1098-1111. doi: 10.1016/j.compedu.2010.05.007

Casey, B., Kersh, J. E., & Young, J. M. (2004). Storytelling sagas: An effective medium for teaching early childhood mathematics. *Early Childhood Research Quarterly, 19*(1), 167-172. doi: 10.1016/j.ecresq.2004.01.011

Cavalier, D. (2014). Science Cheerleader. Retrieved from www.sciencecheerleader.com

Ceci, S. J., & Williams, W. M. (2010). *The mathematics of sex: How biology and society conspire to limit talented women and girls*. New York: Oxford University Press.

Cherney, I. D., Kelly-Vance, L., Gill Glover, K., Ruane, A., & Oliver Ryalls, B. (2003). The effects of stereotyped toys and gender on play assessment in children aged 18-47 months. *Educational Psychology, 23*(1), 95-106. doi: 10.1080/01443410303222

Cheryan, S., Meltzoff, A. N., & Kim, S. (2011). Classrooms matter: The design of virtual classrooms influences gender disparities in computer science classes. *Computers & Education, 57*(2), 1825-1835. doi: 10.1016/j.compedu.2011.02.004

Cheryan, S., Plaut, V. C., Davies, P. G., & Steele, C. M. (2009). Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology, 97*(6), 1045-1060. doi: 10.1037/a0016239

Cowan, C. P., & Cowan, P. A. (1990). Who does what? In J. Touliatos, B. F. Perlmutter & M. A. Straus (Eds.), *Handbook of Family Measurement Techniques* (pp. 447-448). Beverly Hills, CA: Sage.

Coyle, E. F., & Fulcher, M. (2011, March). *Toys as models: How Barbie may impact girls' possible self content*. Poster presented at the Society For Research in Child Development, Montreal, Quebec.

Coyle, E. F., & Liben, L. S. (2013, April). *Attracting girls to STEM: Personal gender salience moderates the outcome of an occupational game*. Poster presented at the Society for Research in Child Development, Seattle, WA.

De Lisi, R., & Wolford, J. L. (2002). Improving children's mental rotation accuracy with computer game playing. *The Journal of Genetic Psychology: Research and Theory on Human Development, 163*(3), 272-282. doi: 10.1080/00221320209598683

Degner, J., & Dalege, J. (2013). The apple does not fall far from the tree, or does it? A meta-analysis of parent-child similarity in intergroup attitudes. *Psychological Bulletin, 139*(6), 1270-1304. doi: 10.1037/a0031436

Eccles, J. S. (2014). Gender and achievement choices. In E. T. Gershoff, R. S. Mistry & D. A. Crosby (Eds.), *Societal contexts of child development: Pathways of influence and implications for practice and policy* (pp. 19-34). New York: Oxford University Press.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75-146). San Francisco: Freeman.

Expanding Your Horizons Network. (2013). Come to an EYH conference: What happens at an EYH conference. Retrieved from http://www.eyhn.org/

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. doi: 10.3758/BF03193146

Fulcher, M. (2011). Individual differences in children’s occupational aspirations as a function of parental traditionality. *Sex Roles*, *64*(1), 117-131. doi: 10.1007/s11199-010-9854-7

Fulcher, M., & Coyle, E. F. (2011). Breadwinner and caregiver: A cross-sectional analysis of children's and emerging adults' visions of their future family roles. *British Journal of Developmental Psychology, 29*(2), 330-346. doi: 10.1111/j.2044-835X.2011.02026.x

Fulcher, M., Sutfin, E. L., & Patterson, C. J. (2008). Individual differences in gender development: Associations with parental sexual orientation, attitudes, and division of labor. *Sex Roles, 58*(5), 330-341. doi: 10.1007/s11199-007-9348-4

Fulcher, M., Weisgram, E. S., & Dinella, L. M. (2013, April). *The effects of implicit and explicit gender labels on children’s interest in novel toys*. Poster presented at the Society for Research in Child Development, Seattle, WA.

Halim, M. L., Ruble, D., Tamis-LeMonda, C., & Shrout, P. E. (2013). Rigidity in gender-typed behaviors in early childhood: A longitudinal study of ethnic minority children. *Child Development*, *84*(4), 1269-1284. doi: 10.1111/cdev.12057

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2008). Sex, math and scientific achievement. *Scientific American*, *18*, 44-51. doi: 10.1038/scientificamericanmind1207-44

Harris, J. R. (1995). Where is the child's environment? A group socialization theory of development. *Psychological Review, 102*(3), 458-489. doi: 10.1037/0033-295X.102.3.458

Harris, J. R. (2009). *The nurture assumption: Why children turn out the way they do*. New York: Free Press.

Helwig, A. A. (1998). Gender-role stereotyping: Testing theory with a longitudinal sample. *Sex Roles, 38*(5-6), 403-423. doi: 10.1023/A:1018757821850

Jirout, J. J., & Newcombe, N. S. (2015). Building blocks for developing spatial skills: Evidence from a large, representative U.S. sample. *Psychological Science,* Advance online publication. doi: 10.1177/0956797614563338

Kelleher, C., & Pausch, R. (2007). Using storytelling to motivate programming. *Communications of the ACM, 50*(7), 58-64. doi: 10.1145/1272516.1272540

Kelleher, C., Pausch, R., & Kiesler, S. (2007). Storytelling Alice motivates middle school girls to learn computer programming. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1455-1464.

Kirk, R. E. (2013). *Experimental design: Procedures for the behavioral sciences* (4th ed.). Thousand Oaks, CA: SAGE Publications.

Levine, S. C., Ratliff, K. R., Huttenlocher, J., & Cannon, J. (2012). Early puzzle play: A predictor of preschoolers' spatial transformation skill. *Developmental Psychology, 48*(2), 530-542. doi: 10.1037/a0025913

Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues, 2*(4), 34-46.

Liben, L. S. (1995). Educational applications of geographic information systems: A developmental psychologist's perspective. In D. Barstow (Ed.), *First national conference on the educational applications of geographic information systems* (pp. 44-49). Cambridge, MA: TERC.

Liben, L. S. (2012, April). *Making STEM girly or making girls STEMy? Rationales and ramifications*. Paper presented at the Gender Development Research Conference, San Francisco, CA.

Liben, L. S., & Bigler, R. (2002). The developmental course of gender differentiation: Conceptualizing, measuring, and evaluating constructs and pathways. *Monographs of the Society for Research in Child Development, 67*(2).

Liben, L. S., Bigler, R. S., & Krogh, H. R. (2001). Pink and blue collar jobs: Children's judgments of job status and job aspirations in relation to sex of worker. *Journal of Experimental Child Psychology*, *79*(4), 346-363. doi: 10.1006/jecp.2000.2611

Liben, L. S., & Coyle, E. F. (2014). Developmental interventions to address the STEM gender gap: Exploring intended and unintended consequences. In L. S. Liben & R. S. Bigler (Eds.), *The role of gender in educational contexts and outcomes* (Vol. 47, pp. 77-115). San Diego, CA: Elsevier.

Liben, L. S., & Downs, R. M. (1986). *Children's production and comprehension of maps: Increasing graphic literacy*. (#G-83-0025). Final Report to National Institute of Education.

Liben, L. S., & Hilliard, L. J. (2010, October). *Preschoolers' gender vigilance: Effects of classroom organization*. Poster presented at the Gender Development Research Conference, San Francisco.

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin*, *136*(6), 1123-1135. doi: 10.1037/a0021276

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56*(6), 1479-1498.

LoBue, V., & DeLoache, J. S. (2011). Pretty in pink: The early development of gender-stereotyped colour preferences. *British Journal of Developmental Psychology, 29*(3), 656-667. doi: 10.1111/j.2044-835X.2011.02027.x

Lynn, K.-M., Raphael, C., Olefsky, K., & Bachen, C. M. (2003). Bridging the gender gap in computing: An integrative approach to content design for girls. *Journal of Educational Computing Research, 28*(2), 143-162. doi: 10.2190/79HP-RVE7-3A9N-FV8C

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

Mammes, I. (2004). Promoting girls' interest in technology through technology education: A research study. *International Journal of Technology and Design Education, 14*(2), 89-100. doi: 10.1023/B:ITDE.0000026472.27439.f6

Mangold, P. (2013). Interact (Version 9.8) [Computer software]. Amstorf, Germany: Mangold International.

Martin, C. L., Eisenbud, L., & Rose, H. (1995). Children's gender-based reasoning about toys. *Child Development, 66*(5), 1453-1471. doi: 10.1111/j.1467-8624.1995.tb00945.x

Martin, C. L., Fabes, R. A., Hanish, L., Leonard, S., & Dinella, L. M. (2011). Experienced and expected similarity to same-gender peers: Moving toward a comprehensive model of gender segregation. *Sex Roles, 65*(5), 421-434. doi: 10.1007/s11199-011-0029-y

Martin, C. L., Fabes, R. A., Hanish, L. D., & Hollenstein, T. (2005). Social dynamics in the preschool. *Developmental Review*, *25*(3), 299-327. doi: 10.1016/j.dr.2005.10.001

Martin, C. L., & Halverson, C. F. (1981). A schematic processing model of sex typing and stereotyping in children. *Child Development, 42*(4), 1119-1134. doi: 10.2307/1129498

Martin, C. L., Kornienko, O., Schaefer, D. R., Hanish, L. D., Fabes, R. A., & Goble, P. (2012). The role of sex of peers and gender-typed activities in young children's peer affiliative networks: A longitudinal analysis of selection and influence. *Child Development, 84*(3), 921-937. doi: 10.1111/cdev.12032

Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010). Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science, 330*(6008), 1234-1237. doi: 10.1126/science.1195996

Molloy, L. E., Gest, S. D., & Rulison, K. L. (2011). Peer influences on academic motivation: Exploring multiple methods of assessing youths' most "influential" peer relationships. *The Journal of Early Adolescence, 31*(1), 13-40. doi: 10.1177/0272431610384487

Murphy, P., & Whitelegg, E. (2006). Girls and physics: Continuing barriers to "belonging". *Curriculum Journal, 17*(3), 281-305. doi: 10.1080/09585170600909753

Parker, K., & Wang, W. (2013). Modern parenthood: Roles of moms and dads converge as they balance work and family. *Pew Research Center: Social and Demographic Trends*. Retrieved from http://www.pewsocialtrends.org/2013/03/14/modern-parenthood-roles-of-moms-and-dads-converge-as-they-balance-work-and-family/

Patterson, M. M. (2012). Self-perceived gender typicality, gender-typed attributes, and gender stereotype endorsement in elementary-school-aged children. *Sex Roles, 67*(7-8), 422-434. doi: 10.1007/s11199-012-0184-9

Piaget, J., & Inhelder, B. (1956). *The child's conception of space*. London: Routledge & Kegan Paul Ltd.

Rodger, S. H., Hayes, J., Lezin, G., Qin, H., Nelson, D., Tucker, R., . . . Slater, D. (2009). Engaging middle school teachers and students with Alice in a diverse set of subjects. *ACM SIGCSE Bulletin, 41*(1), 271-275.

Rubin, K. H., Bukowski, W. M., & Parker, J. G. (2006). Peer interactions, relationships, and groups. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology* (Vol. 3). Hoboken, NJ: Wiley.

Serbin, L. A., & Sprafkin, C. (1986). The salience of gender and the process of sex typing in three-to seven-year-old children. *Child Development, 57*(5), 1188-1199. doi: 10.2307/1130442

Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology, 93*(3), 604. doi: 10.1037/0022-0663.93.3

Simpkins, S. D., Fredricks, J. A., & Eccles, J. S. (2012). Charting the Eccles' expectancy-value model from mothers' beliefs in childhood to youths' activities in adolescence. *Developmental Psychology, 48*(4), 1019-1032. doi: 10.1037/a0027468

Smetana, J. G. (1986). Preschool children's conceptions of sex-role transgressions. *Child Development, 57*(4), 862-871.

Sorby, S., Casey, B., Veurink, N., & Dulaney, A. (2013). The role of spatial training in improving spatial and calculus performance in engineering students. *Learning and Individual Differences*. doi: 10.1016/j.lindif.2013.03.010

Sweet, E. (2014). Toys are more divided by gender now than they were 50 years ago. Retrieved from http://www.theatlantic.com/business/archive/2014/12/toys-are-more-divided-by-gender-now-than-they-were-50-years-ago/383556/

Swim, J. K., Aikin, K. J., Hall, W. S., & Hunter, B. A. (1995). Sexism and racism: old-fashioned and modern prejudices. *Journal of Personality and Social Psychology, 68*(2), 199-214. doi: 10.1037/0022-3514.68.2.199

Taillandier-Schmitt, A., Esnard, C., & Mokounkolo, R. (2012). Self-affirmation in occupational training: Effects on the math performance of French women nurses under stereotype threat. *Sex Roles, 67*(1-2), 43-57. doi: 10.1007/s11199-012-0157-z

Tenenbaum, H. R., & Leaper, C. (2002). Are parents' gender schemas related to their children's gender-related cognitions? A meta-analysis. *Developmental Psychology, 38*(4), 615-630. doi: 10.1037//0012-1649.38.4.615

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin, 139*(2), 352-402. doi: 10.1037/a0028446

Wai, J., Cacchio, M., Putallaz, M., & Makel, M. C. (2010). Sex differences in the right tail of cognitive abilities: A 30 year examination. *Intelligence, 38*(4), 412-423. doi: 10.1016/j.intell.2010.04.006

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101*(4), 817-835. doi: 10.1037/a0016127

Weisgram, E. S., & Bigler, R. S. (2006). Girls and science careers: The role of altruistic values and attitudes about scientific tasks. *Journal of Applied Developmental Psychology, 27*(4), 326-348. doi: 10.1016/j.appdev.2006.04.004

Weisgram, E. S., & Bigler, R. S. (2007). Effects of learning about gender discrimination on adolescent girls' attitudes toward and interest in science. *Psychology of Women Quarterly, 31*(3), 262-269. doi: 10.1111/j.1471-6402.2007.00369.x

Wolfgang, C., Stannard, L., & Jones, I. (2003). Advanced constructional play with LEGOs among preschoolers as a predictor of later school achievement in mathematics. *Early Child Development and Care, 173*(5), 467-475. doi: 10.1080/0300443032000088212

# APPENDIX A

Tables and Figures

Table 1

*Family Characteristics*

| | Min | Max | Mean | SD | Median | Mode |
|---|---|---|---|---|---|---|
| *Mother* | | | | | | |
| Age | 25.00 | 62.00 | 36.13 | 5.82 | 35.50 | 32.00 |
| Education | 2.00 (High school diploma/G.E.D.) | 6.00 (Graduate degree) | 5.07 (Bachelor's degree) | 1.05 | 5.00 | 5.00 |
| Job Traditionality | 1.00 | 3.57 | 2.20 | 0.91 | 2.54 | 1.00 |
| Job Prestige | 1.15 | 3.22 | 1.96 | 0.68 | 2.13 | 1.15 |
| Hours Employed | 0.00 (Stay at home) | 3.00 (Full-Time) | 1.58 | 1.24 | 2.00 | 2.00 |
| *Father* | | | | | | |
| Age | 26.00 | 60.00 | 37.76 | 6.64 | 36.00 | 33 |
| Education | 1.00 (Some high school) | 6.00 | 4.67 | 1.23 | 5.00 | 5.00 |
| Job Traditionality | 1.00 | 4.45 | 3.31 | 0.56 | 3.28 | 3.28 |
| Job Prestige | 1.15 | 3.65 | 2.61 | 0.52 | 2.58 | 3.22 |
| Hours Employed | 0.00 | 3.00 | 2.87 | 0.47 | 3.00 | 3.00 |

*Household Income*

| | | | | |
|---|---|---|---|---|
| 18% $0-$50,000 | 26% $50-80,000 | 38% $80-120,000 | 16% $120,000+ | 2% declined to answer |

*Race/ethnicity*

| | | | | |
|---|---|---|---|---|
| Mother | 93% White | 2% Black | 3% Asian | 2% Latina |
| Father | 92% White | 2% Black | 3% Asian | 3% Latino |

Table 2

*Procedure*

| Activity | Target |
| --- | --- |
| *Experimental Session* | |
| Familiarize with the toy (5 minutes) | Parent |
| During Parent familiarization:<br>Survey about favorite toys,<br>Water-level task | Child<br>(with Researcher) |
| Dyadic play (15 minutes) | Parent and Child |
| Individual play (7 minutes) | Child |
| Mechanical test (5 problems)<br>Transfer tasks | Child<br>(with Researcher) |
| Interest in *Goldie/BobbyBlox*<br>*Goldie/BobbyBlox* Character Typicality<br>Interest in playing with STEM toys in the future | Child<br>(with Researcher) |
| During Child Individual play and Mechanical testing:<br>Demographic survey<br>Interest in *Goldie/BobbyBlox*<br>Interest in purchasing other STEM toys | Parent |
| *Online Session, 5-7 days after the experimental session* | |
| Gender measures<br>OAT-PM Activities and Traits<br>Who Does What?<br>Feminism/Collective Action<br>Parent-Child Expectancy Value<br>COAT Occupations (for Child)<br>Gender Socialization Scale<br>OAT-AM<br>Modern Sexism | Parent |
| Gender measures<br>POAT-PM Occupations and Activities<br>Implicit Memory for Gender – Encoding Phase<br>Gendered Affiliation<br>Implicit Memory for Gender<br>Gender Typicality<br>POAT-AM Occupations and Activities | Child |

Table 3

*Masculinity-Femininity Ratings of Game Characters*

| Character | Min | Max | Mean | SD |
|---|---|---|---|---|
| Flavio | 1.00 | 3.00 | 1.54 | 0.63 |
| Phil | 1.00 | 4.00 | 2.29 | 0.74 |
| Nacho | 1.00 | 5.00 | 2.49 | 0.86 |
| Benjamin | 1.00 | 5.00 | 3.16 | 0.81 |
| Katina | 1.00 | 5.00 | 4.93 | 0.29 |

*Note*. Ratings were made by Psychology Subject Pool participants on a scale from 1(very masculine) to 5 (very feminine).

Table 4

*Descriptive Statistics for Spatial and Mechanical Measures*

| Measure | Min | Max | *M* (*SD*) Girls | *M* (*SD*) Boys |
|---|---|---|---|---|
| WLT | 0.00 | 6.00 | 2.55 (1.36) | 2.03 (1.35) |
| Mechanical Test | 0.00 | 5.00 | 2.39 (1.20) | 2.13 (1.50) |
| Easy Transfer Task | 57.00 | 509.00 | 222.65 (108.85) | 188.60 (92.21) |
| Difficult Transfer Task | 1.00 | 4.00 | 2.40 (1.28) | 2.38 (1.29) |
| | | *Mothers* | | |
| WLT (5 degrees) | 0.00 | 6.00 | 3.80 (1.95) | |
| WLT (10 degrees) | 0.00 | 6.00 | 4.95 (1.62) | |

Table 5

*Descriptive Statistics for Mothers' Game Interest Outcome Measures*

| | Min | Max | *M* (*SD*) Goldie/ Daughters | *M* (*SD*) Goldie/Sons | *M* (*SD*) Bobby/ Daughters | *M* (*SD*) Bobby/Sons |
|---|---|---|---|---|---|---|
| *Goldie/BobbyBlox* | | | | | | |
| Enjoyment | 2.00 | 4.00 | 3.27 (0.46) | 3.13 (0.64) | 3.44 (0.63) | 3.33 (0.49) |
| Interest in Buying | 1.00 | 4.00 | 2.79 (0.70) | 2.47 (0.83) | 2.81 (0.66) | 2.80 (0.86) |
| Educational Value | 2.00 | 4.00 | 3.40 (0.51) | 3.40 (0.63) | 3.69 (0.48) | 3.47 (0.52) |
| *Interest in Buying STEM Toys* | | | *M* (*SD*) Daughters | | *M* (*SD*) Sons | |
| Feminine | 1.20 | 3.80 | 2.80 (0.50 | | 2.36 (0.44) | |
| Neutral | 2.20 | 4.00 | 3.08 (0.49) | | 3.11 (0.43) | |
| Masculine | 1.00 | 3.80 | 2.54 (0.62) | | 3.08 (0.44) | |

Table 6

*Descriptive Statistics for Children's Game Interest Outcome Measures*

| | Min | Max | *M* (*SD*) Goldie/Girls | *M* (*SD*) Goldie/Boys | *M* (*SD*) Bobby/Girls | *M* (*SD*) Bobby/Boys |
|---|---|---|---|---|---|---|
| *Interest in Goldie/BobbyBlox* | | | | | | |
| Overall | 1.00 | 3.00 | 2.80 (0.41) | 2.13 (0.64) | 2.63 (0.50) | 2.33 (0.72) |
| At Home | 1.00 | 3.00 | 2.67 (0.49) | 2.27 (0.80) | 2.44 (0.73) | 1.87 (0.83) |
| At School | 1.00 | 3.00 | 2.73 (0.46) | 1.86 (0.86) | 2.38 (0.81) | 2.13 (0.83) |
| Similarity to Character | 1.00 | 3.00 | 2.22 (0.51) | 1.72 (0.54) | 2.00 (0.28) | 1.71 (0.63) |
| *Interest in STEM Toys* | | | *M* (*SD*) All Girls | | *M* (*SD*) All Boys | |
| Feminine | 1.00 | 3.00 | 2.73 (0.29) | | 1.91 (0.53) | |
| Neutral | 1.00 | 3.00 | 2.26 (0.53) | | 2.43 (0.47) | |
| Masculine | 1.20 | 3.00 | 2.03 (0.58) | | 2.53 (0.39) | |

Table 7

*Descriptive Statistics for Mothers' Individual Difference Measures*

| Measure (Possible Range) | Min | Max | *M* (*SD*) |
|---|---|---|---|
| *Interests* | | | |
| OAT-PM Activities Feminine (1-4) | 2.60 | 3.78 | 3.20 (0.26) |
| OAT-PM Activities Masculine (1-4) | 1.20 | 2.40 | 1.74 (0.29) |
| OAT-PM Traits Feminine (1-4) | 1.90 | 3.50 | 2.79 (0.34) |
| OAT-PM Traits Masculine (1-4) | 1.80 | 3.50 | 2.76 (0.39) |
| *Attitudes* | | | |
| OAT-AM Activities (Flexibility; 0-1) | .45 | 1.00 | .88 (.16) |
| Collective Action (Feminism; 1-4) | 1.20 | 4.00 | 2.23 (0.63) |
| Modern Sexism (1-4) | 1.14 | 3.00 | 2.05 (0.36) |
| *Division of Labor* | | | |
| Who Does What: Household Labor (1-9) | 4.58 | 9.00 | 6.05 (1.12) |
| Who Does What: Decision-making (1-9) | 4.17 | 9.00 | 5.71 (1.06) |
| Who Does What: Childcare (1-9) | 3.00 | 9.00 | 6.89 (1.13) |

Table 8

*Descriptive Statistics for Mothers' Beliefs about their Children*

| Measure (Possible Range) | Min | Max | *M* (*SD*) Daughters | *M* (*SD*) Sons |
|---|---|---|---|---|
| *Parent-Child Expectancy-Value* | | | | |
| Sports Ability | 2.33 | 7.00 | 4.73 (0.80) | 4.97 (1.10) |
| Music Ability (1-7) | 3.33 | 7.00 | 4.75 (0.79) | 4.80 (1.14) |
| Math Ability (1-7) | 3.33 | 7.00 | 5.13 (1.00) | 5.47 (0.89) |
| Reading Ability (1-7) | 3.00 | 7.00 | 5.15 (1.12) | 4.87 (1.08) |
| Importance of Sports (1-7) | 1.00 | 7.00 | 4.38 (1.53) | 3.60 (1.54) |
| Importance of Music (1-7) | 1.00 | 7.00 | 4.75 (1.29) | 4.15 (1.69) |
| Importance of Math (1-7) | 3.00 | 7.00 | 5.75 (1.26) | 6.40 (0.75) |
| Importance of Reading (1-7) | 3.00 | 7.00 | 6.17 (1.31) | 6.60 (0.68) |
| COAT-PM Feminine Jobs (1-7) | 1.10 | 6.30 | 4.64 (0.78) | 3.83 (1.27) |
| COAT-PM Masculine Jobs (1-7) | 3.50 | 7.00 | 5.31 (0.73) | 5.42 (0.76) |
| *Gender Socialization* | | | | |
| Girl activities (1-7) | 3.88 | 7.00 | 5.76 (0.92) | 5.43 (0.94) |
| Boy activities (1-7) | 3.29 | 6.71 | 4.62 (0.84) | 5.12 (0.98) |
| Helping at home (1-7) | 3.86 | 7.00 | 6.35 (0.84) | 6.56 (0.64) |
| Education for future family (1-7) | 1.00 | 7.00 | 4.83 (1.75) | 6.12 (1.05) |
| Education for future career (1-7) | 5.00 | 7.00 | 6.60 (0.57) | 6.92 (0.18) |
| Discourage nonconformity (1-7) | 1.00 | 6.00 | 2.08 (1.24) | 2.30 (1.40) |

Table 9

*Descriptive Statistics for Children's Individual Difference Measures*

| Measure (Possible Range) | Min | Max | *M* (*SD*) Girls | *M* (*SD*) Boys |
|---|---|---|---|---|
| *Interests* | | | | |
| POAT-PM Feminine Jobs (1-3) | 1.00 | 3.00 | 2.46 (0.28) | 1.65 (0.30) |
| POAT-PM Masculine Jobs (1-3) | 1.00 | 3.00 | 1.92 (0.62) | 2.37 (.44) |
| POAT-PM Feminine Activities (1-3) | 1.00 | 3.00 | 2.67 (0.53) | 1.61 (0.44) |
| POAT-PM Masculine Activities (1-3) | 1.00 | 3.00 | 2.17 (0.50) | 2.59 (0.28) |
| *Attitudes* | | | | |
| POAT-AM Occupations (1-3) | .08 | 1.00 | .49 (.23) | .57 (.57) |
| POAT-AM Activities | .17 | 1.00 | .52 (.27) | .58 (.24) |
| *Gender Salience* | | | | |
| Composite GSF Score | -3.24 | 2.57 | 0.40 (1.47) | -.32 (1.38) |
| Gender Memory (1-3) | .00 | .90 | .72 (.20) | .69 (.14) |
| Girl Typicality (1-3) | 1.00 | 3.00 | 2.88 (0.26) | 1.39 (0.35) |
| Boy Typicality (1-3) | 1.00 | 3.00 | 1.65 (0.47) | 2.69 (0.35) |
| Affiliation (0-1) | .00 | 1.00 | .64 (.31) | .60 (.25) |
| Feminine Clothing (0-8, girls only) | 0.00 | 6.00 | 2.90 (1.62) | -- |
| Masculine Clothing (0-5, boys only) | 2.00 | 4.00 | -- | 3.00 (0.53) |
| Favorite Toy (1-5) | 1.22 | 4.49 | 2.14 (0.73) | 3.53 (0.55) |

Table 10

*Mother Familiarization Video Codes*

| Code | Definition | Example |
|---|---|---|
| Box | Examining box/packaging | Mother reads the back of the GoldieBlox box |
| Book | Reading book | Mother reads the BobbyBlox book |
| Pieces | Examining the pieces | Mother counts the wheels in the GoldieBlox set |
| Building | Building with the set | Mother puts an axle in the pegboard and adds a wheel |
| Other | Anything else | Mother checks her cellphone |

*Note.* Videos were coded in 10-second intervals with one code per interval. Within an interval, the majority behavior was coded for the interval.

Kappa = .97.

Table 11

*Dyadic Play Behaviors*

| Code | Definition | Example |
|---|---|---|
| Book Building | Attempting to build from the book | Dyad builds the two-wheel belt drive from the book |
| Other Building | Building unrelated to the book | Dyad builds a tower with the pieces in the set |
| Reading | Reading the story (no building) | Dyad reads the story but is not using any pieces in the set for building |
| Pieces | Counting or touching the pieces | Dyad is counting the number of wheels in the set |
| Other | Anything else | Dyad is discussing what "engineering" means |

*Note.* Videos were coded in 8-second intervals with one code per interval. Within an interval, the majority behavior was coded for the interval.

Kappa = .78.

Table 12

*Mother Scaffolding in Dyadic Play*

| Code | Definition | Example |
|---|---|---|
| Demonstrate | Demonstrate building | Mother demonstrates how to wind the ribbon around the wheel to create the belt-drive |
| Instruct | Instruct how to build | Mother tells the child to put the axle in the pegboard |
| Correct | Correct a building error | Child attaches the ribbon Velcro at both ends and mother detaches or tells the child to detach |
| Read | Read the book | Mother reads the book for building instructions |
| Label | Label a piece or structure | Mother picks up axle and says “this is an axle” |

*Note.* Videos were coded in 10-second intervals, allowing multiple codes per interval.

Kappa = .74.

Table 13

*Child Free Play Behaviors*

| Code | Definition | Example |
|---|---|---|
| Book Building | Attempting to build from the book | Child attempts to build a design from the back pages of the book |
| Other Building | Building unrelated to the book | Child builds a tower with the pieces in the set |
| Pretend Play | Reading the story (no building) | Child narrates as the animals in the set |
| Not Playing | Anything without the set | Child stops playing and walks around the room |

*Note.* Videos were coded in 5-second intervals with one code per interval. Within an interval, the majority behavior was coded for the interval.

Kappa = .91.

Table 14

*Correlations among Children's Key Study Variables within the GoldieBlox Condition*

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. GSF | - | -.16 | .39 | .10 | .10 | .21 | .37 | -.34 | .12 | .04 | -.13 | -.63* | -.65* |
| 2. Favorite Toy | .08 | - | .24 | .06 | .23 | .31 | .22 | .25 | -.26 | -.46 | .19 | .20 | .04 |
| 3. Blox Interest | -.49 | .11 | - | .35 | .83** | .57* | .26 | .29 | -.61* | -.49 | .29 | -.09 | .10 |
| 4. Blox: Home | .43 | .38 | .48 | - | .21 | .38 | .06 | .02 | .02 | -.18 | .82** | .27 | .22 |
| 5. Blox: School | -.64 | -.31 | .31 | .27 | - | .28 | .25 | .34 | -.74** | -.60* | .21 | .11 | .29 |
| 6. Blox: Similarity | -.42 | .06 | .57* | .22 | .35 | - | .03 | .39 | -.44 | -.43 | .23 | -.05 | -.02 |
| 7. Mechanical Score | -.12 | -.16 | .08 | .28 | .35 | .03 | - | -.37 | .07 | .02 | .20 | -.31 | -.48 |
| 8. Easy Transfer | .71 | .05 | -.03 | .24 | -.25 | -.01 | -.35 | - | -.67** | -.21 | -.21 | .04 | .13 |
| 9. Difficult Transfer | -.10 | -.30 | -.01 | -.09 | -.01 | .17 | .20 | -.23 | - | .69** | .00 | -.07 | -.35 |
| 10. WLT | -.12 | -.13 | .20 | .00 | -.32 | .28 | -.20 | .02 | .74** | - | -.37 | -.19 | -.33 |
| 11. STEM Toys F | -.57 | -.03 | .13 | -.18 | .39 | .21 | .10 | -.05 | .03 | -.20 | - | .47 | .35 |
| 12. STEM Toys N | -.07 | .53* | .51 | .75** | .21 | .01 | .24 | .11 | -.19 | -.20 | -.02 | - | .82** |
| 13. STEM Toys M | -.63 | .56* | .29 | .17 | .19 | .19 | .13 | -.65** | -.20 | -.23 | .14 | .40 | - |

*Note*. Correlations for girls are given above the diagonal. Correlations for boys are given below the diagonal. Variable 1 refers to children's composite gender salience filter (GSF) score, calculated as the sum of the standardized scores for gender memory, gendered affiliation, and gender typicality. Variable 2 refers to the gendered rating of children's favorite toys. Variables 3-5 refer to children's interest in *GoldieBlox*. Variable 6 refers to children's perceived similarity to Goldie. Variables 7-9 are children's scores on the mechanical and transfer outcome measures. Variable 10 refers to the water-level task assessment of spatial ability. Variables 11-13 refer to children's interest in future play with feminine (F), neutral (N), and masculine (M) STEM toys.
* $p < .05$. ** $p < .01$.

Table 15

*Correlations among Children's Key Study Variables within the BobbyBlox Condition*

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. GSF | - | -.30 | .02 | .41 | .25 | -.11 | .45 | .38 | .50 | -.11 | .11 | .57 | .51 |
| 2. Favorite Toy | -.28 | - | -.14 | .09 | .29 | .28 | -.38 | .12 | .14 | .54* | .10 | .46 | .34 |
| 3. Blox Interest | -.07 | .49 | - | .12 | .37 | .00 | .01 | .54* | .14 | -.04 | -.05 | .08 | .01 |
| 4. Blox: Home | .26 | .22 | .55* | - | .50* | .63** | .46 | .16 | .18 | -.14 | -.08 | .51 | .38 |
| 5. Blox: School | .16 | .63* | .87** | .54* | - | .25 | .11 | .60* | .46 | .14 | .03 | .69** | .45 |
| 6. Blox: Similarity | -.58 | .12 | .11 | -.22 | .22 | - | .11 | .13 | -.09 | .17 | .13 | .55* | .23 |
| 7. Mechanical Score | -.47 | .05 | .43 | .28 | .23 | -.15 | - | -.10 | .15 | .04 | -.57* | -.07 | -.26 |
| 8. Easy Transfer | .23 | -.19 | -.40 | .04 | -.27 | .10 | -.55* | - | .01 | .02 | .21 | .47 | .36 |
| 9. Difficult Transfer | .07 | -.37 | -.29 | -.17 | -.31 | -.21 | .43 | -.46 | - | .42 | -.10 | .40 | .42 |
| 10. WLT | -.15 | .03 | .29 | .39 | .16 | -.14 | .61* | -.66** | .59* | - | -.08 | .27 | .10 |
| 11. STEM Toys F | -.35 | .13 | .53* | .50 | .35 | .16 | .20 | .33 | -.36 | .01 | - | .21 | .45 |
| 12. STEM Toys N | .09 | .52 | .23 | .25 | .31 | .51 | -.23 | .21 | -.31 | -.01 | .11 | - | .64** |
| 13. STEM Toys M | .56 | .55* | .46 | .31 | .65** | .36 | -.11 | -.11 | -.17 | .03 | -.01 | .67** | - |

*Note*. Correlations for girls are given above the diagonal. Correlations for boys are given below the diagonal. Variable 1 refers to children's composite gender salience filter (GSF) score, calculated as the sum of the standardized scores for gender memory, gendered affiliation, and gender typicality. Variable 2 refers to the gendered rating of children's favorite toys. Variables 3-5 refer to children's interest in *BobbyBlox*. Variable 6 refers to children's perceived similarity to Goldie. Variables 7-9 are children's scores on the mechanical and transfer outcome measures. Variable 10 refers to the water-level task assessment of spatial ability. Variables 11-13 refer to children's interest in future play with feminine (F), neutral (N), and masculine (M) STEM toys.

*$p < .05$. **$p < .01$.

Table 16

*Correlations among Mother's Key Study Variables within the GoldieBlox Condition*

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. Work Composite | - | .18 | -.44 | -.21 | -.18 | -.31 | .36 | -.29 | -.23 | .11 |
| 2. Feminism | .20 | - | -.62* | .06 | .32 | -.12 | -.35 | -.07 | -.07 | -.10 |
| 3. Math Importance | .51 | .48 | - | .43 | .37 | .40 | .23 | .38 | .37 | .37 |
| 4. Blox: Enjoy | -.14 | .04 | -.58 | - | .67** | .74** | -.05 | .40 | .49 | .45 |
| 5. Blox: Purchase | .07 | .36 | -.24 | .81** | - | .49 | -.36 | .31 | .65* | .31 |
| 6. Blox: Educational Value | -.05 | .27 | -.43 | .74** | .70** | - | -.25 | .18 | .15 | -.09 |
| 7. WLT (5 degrees) | .03 | .30 | -.11 | .36 | .45 | .46 | - | -.17 | -.14 | .18 |
| 8. STEM Toys F | .32 | .51 | .67* | -.01 | .18 | .17 | .51 | - | .39 | .51 |
| 9. STEM Toys N | -.28 | .79** | .20 | .13 | .13 | .22 | -.12 | .29 | - | .72** |
| 10. STEM Toys M | .18 | .12 | -.17 | .25 | .17 | .45 | .16 | .48 | .30 | - |

*Note*. Correlations for daughters are given above the diagonal. Correlations for sons are given below the diagonal. Variable 1 refers to work traditionality composite score score, calculated as the sum of the standardized scores for occupational femininity, occupational prestige, and hours employed. Variable 2 refers mothers' scores on the Collective Action scale. Variable 3 refers to mothers' ratings math importance for their child. Variables 4-6 refer to mothers' ratings of *GoldieBlox* (enjoyment, interest in purchasing, and educational value). Variable 7 refers to the water-level task assessment of spatial ability (number correct within 5 degrees). Variables 8-10 refer to mothers' interest in purchasing feminine (F), neutral (N), and masculine (M) STEM toys for their child.

*$p < .05$. **$p < .01$.

Table 17

*Correlations among Mother's Key Study Variables within the BobbyBlox Condition*

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. Work Composite | - | .44 | -.14 | .26 | -.02 | .07 | -.62* | -.29 | .00 | -.03 |
| 2. Feminism | .14 | - | .19 | .31 | -.03 | -.42 | -.17 | .06 | .25 | .32 |
| 3. Math Importance | -.33 | -.14 | - | -.40 | .03 | .15 | -.26 | .30 | -.44 | .28 |
| 4. Blox: Enjoy | .31 | -.13 | .30 | - | .21 | .48 | .01 | .43 | -.05 | .06 |
| 5. Blox: Purchase | -.02 | .09 | .32 | .51 | - | .01 | .10 | .01 | .24 | .62* |
| 6. Blox: Educational Value | .02 | -.13 | .61 | .76** | .71** | - | .10 | .37 | -.29 | -.05 |
| 7. WLT (5 degrees) | -.18 | -.04 | -.04 | .19 | .43 | .33 | - | .06 | .34 | .32 |
| 8. STEM Toys F | .37 | -.01 | -.15 | .60* | .30 | .44 | .30 | - | -.20 | -.06 |
| 9. STEM Toys N | .18 | -.50 | .12 | .13 | .23 | .19 | .06 | .17 | - | .42 |
| 10. STEM Toys M | .54* | .15 | -.06 | .43 | .14 | .01 | -.05 | .33 | .26 | - |

*Note*. Correlations for daughters are given above the diagonal. Correlations for sons are given below the diagonal. Variable 1 refers to work traditionality composite score score, calculated as the sum of the standardized scores for occupational femininity, occupational prestige, and hours employed. Variable 2 refers mothers' scores on the Collective Action scale. Variable 3 refers to mothers' ratings math importance for their child. Variables 4-6 refer to mothers' ratings of *BobbyBlox* (enjoyment, interest in purchasing, and educational value). Variable 7 refers to the water-level task assessment of spatial ability (number correct within 5 degrees). Variables 8-10 refer to mothers' interest in purchasing feminine (F), neutral (N), and masculine (M) STEM toys for their child.

**p* < .05. ***p* < .01.

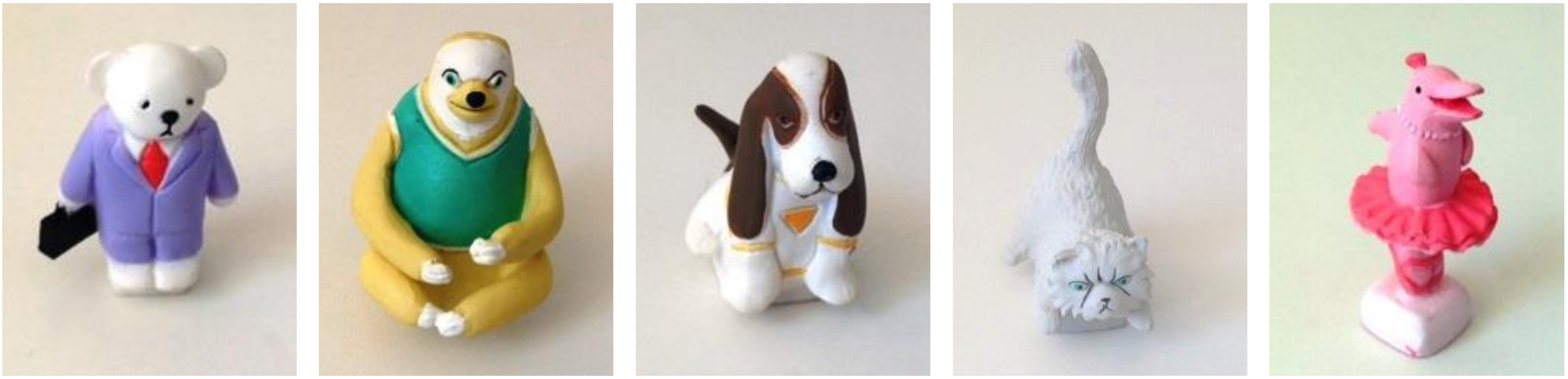
Table 18

*Correlations among Maternal Scaffolding Strategies, Children's Mechanical Learning, and Game Completion*

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1. Mechanical Score | - | .02 | .02 | .16 | .18 | -.23 | -.08 |
| 2. Game Completion | -.06 | - | .55** | -.32 | .39* | -.08 | -.30 |
| 3. Demonstrate Building | -.22 | .25 | - | -.49** | .16 | .08 | -.57** |
| 4. Instruct | -.25 | -.03 | -.69** | - | .32 | -.38* | -.27 |
| 5. Correct Mistake | -.29 | .07 | -.21 | .39* | - | -.23 | -.59** |
| 6. Label Piece/Construction | -.11 | -.08 | -.25 | .24 | .36 | - | -.19 |
| 7. Read Book | .57** | -.28 | -.60** | -.08 | -.42* | -.25 | - |

*Note*. Correlations for daughters are given above the diagonal. Correlations for sons are given below the diagonal. Variable 1 refers to children's score on the mechanical learning task. Variable 2 refers to how far the dyad progressed in the game during dyadic play. Variables 3-7 refer to maternal scaffolding strategies coded from dyadic play.

**p* < .05. ***p* < .01.



*Flavio* *Phil* *Nacho* *Benjamin* *Katinka*

*Figure 1.* Both games included 5 animal figurines. Animals were pre-tested by Psychology Subject Pool participants for masculinity/femininity and fell on a continuum, $F(3.51,347.83) = 367.57$, $p < .001$, $\eta^2_p = .79$. They are depicted from left to right in order from most masculine to most feminine.

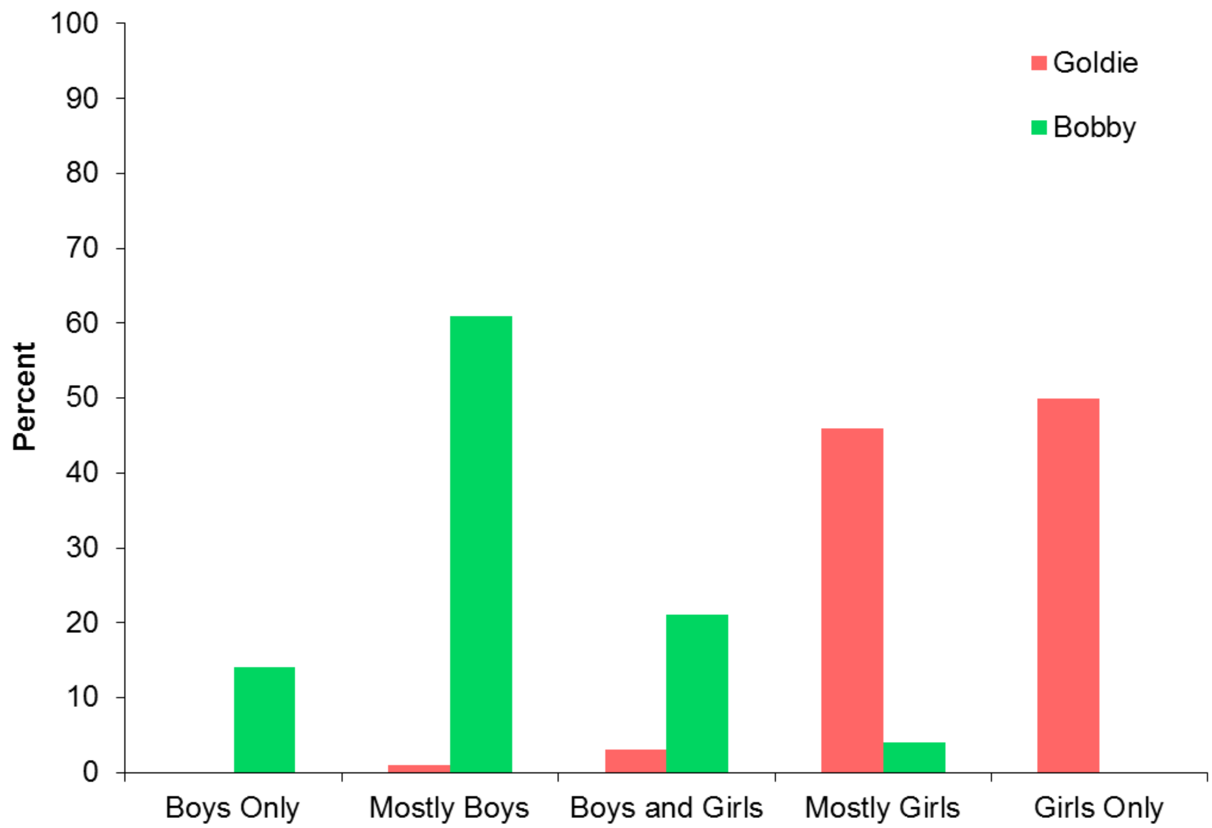



*Figure 2*. Psychology Subject Pool participants rated *BobbyBlox* as a boys' toy and *GoldieBlox* as a girls' toy, confirming the toys were indeed designed to be differentially marketed to boys versus to girls, $F(1,98) = 493.73$, $p < .001$, $\eta^2_p = .83$.

*Box:*


*Pieces:*

*Book:*

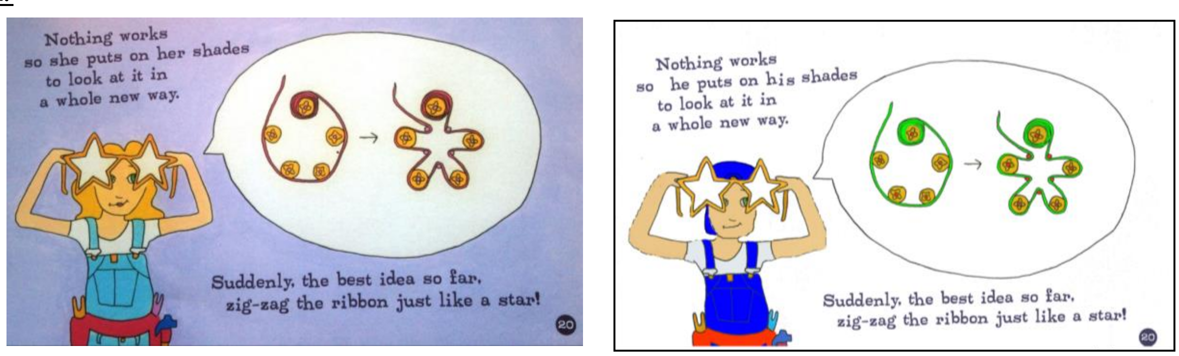



*Figure 3*. Comparison of *GoldieBlox* (left) to *BobbyBlox* (right).

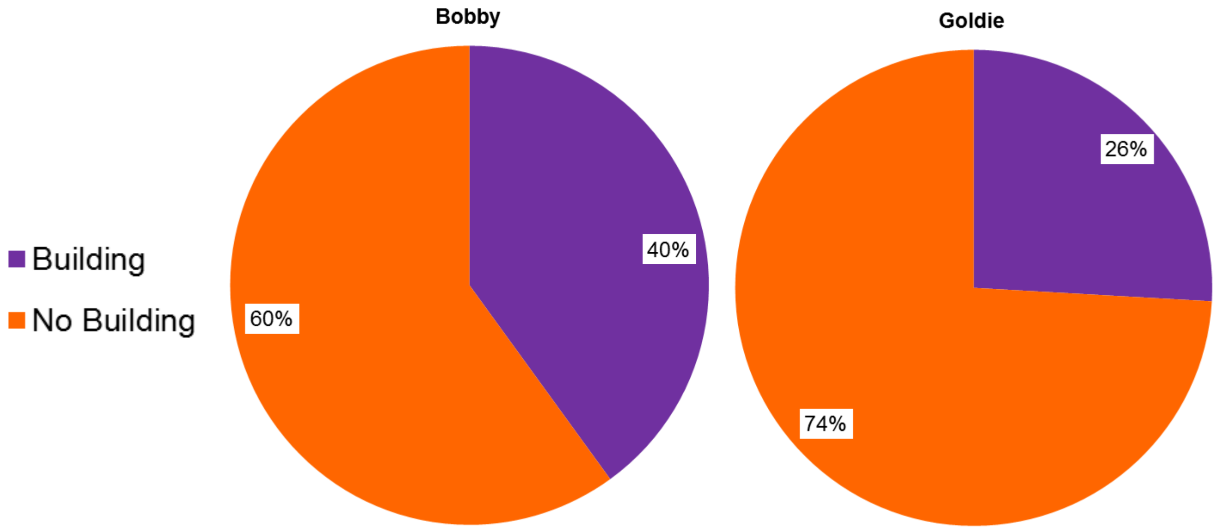



*Figure 4.* Proportion of mothers who engaged in any building during the pre-play familiarization period with *BobbyBlox* versus *GoldieBlox*. Mothers engaged in more building with *BobbyBlox* than with *GoldieBlox*, $F(1,53) = 3.90$, $p = .050$, $\eta^2_p = .07$.

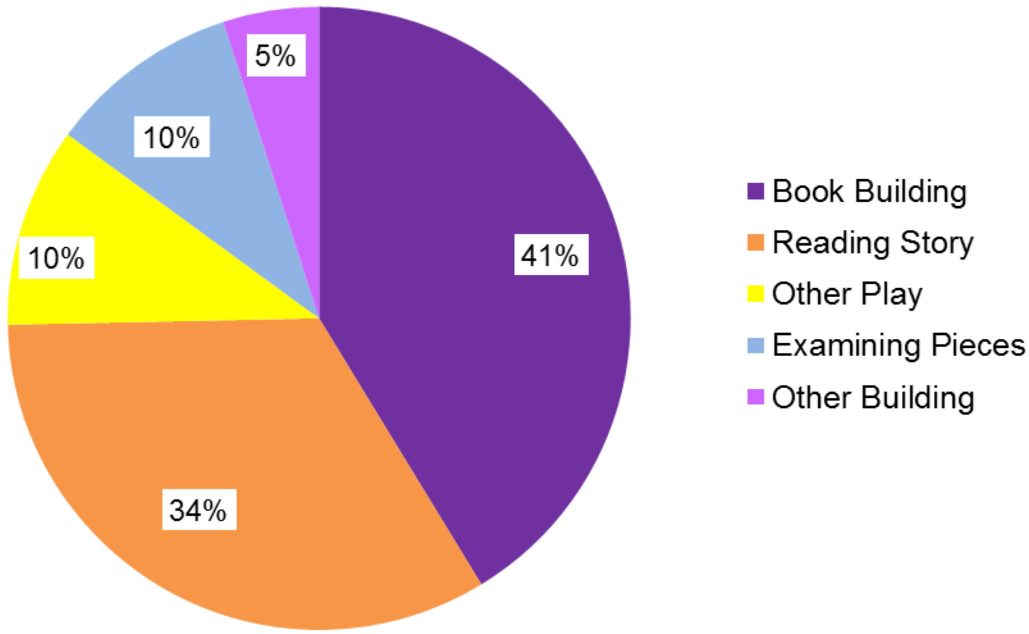



*Figure 5.* Relative frequency of each play behavior code during dyadic play. Dyads spent most of their time reading and building along with the book, $F(2.83,161.34) = 63.61$, $p < .001$, $\eta^2_p = .71$.

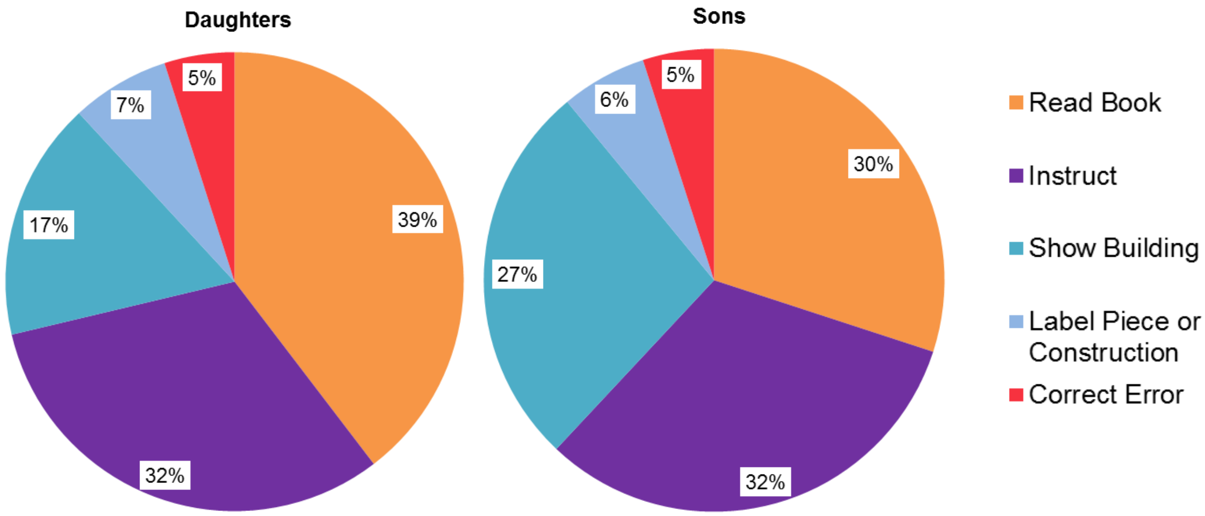



*Figure 6.* Relative frequency of maternal scaffolding strategy use with daughters versus sons. Mothers used instruction and reading most commonly, building demonstration less commonly, and corrected mistakes and labelled pieces least often, $F(2.22, 126.73) = 80.39$, $p < .001$, $\eta^2_p = .59$.

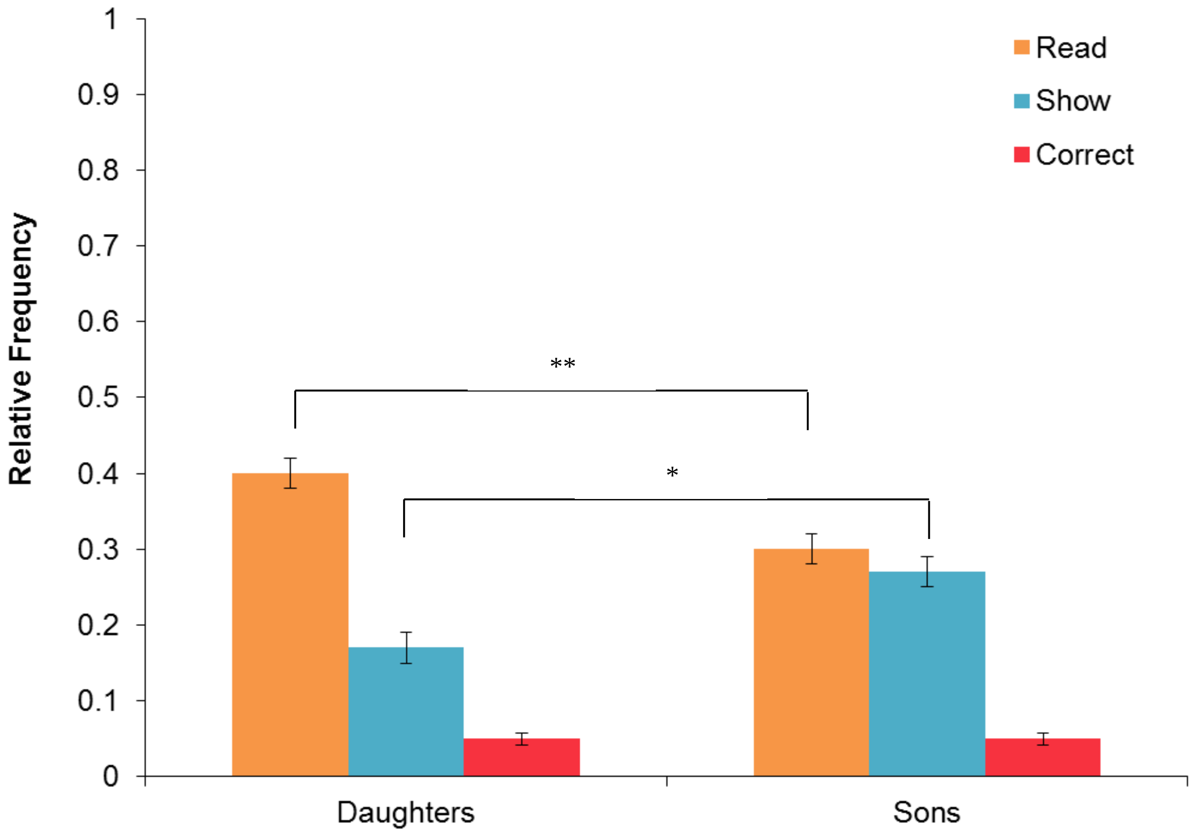



*Figure 7*. Mothers used different scaffolding strategies with daughters versus with sons. They read more to daughters and demonstrated building more to sons, $F(2.22,126.73) = 5.08$, $p = .006$, $\eta^2_p = .03$. Reading the book was the only strategy to predict mechanical learning, $F(1,59) = 5.73$, $p = .020$, $R^2 = .09$, $\beta = 0.30$, $p = .020$. Demonstrating building and correcting mistakes predicted game completion, $F(2,56) = 6.93$, $p = .002$, $R^2 = .20$, $\beta = 0.38$, $p = .002$ (demonstrating building), $\beta = 0.24$, $p = .046$ (correcting mistakes).

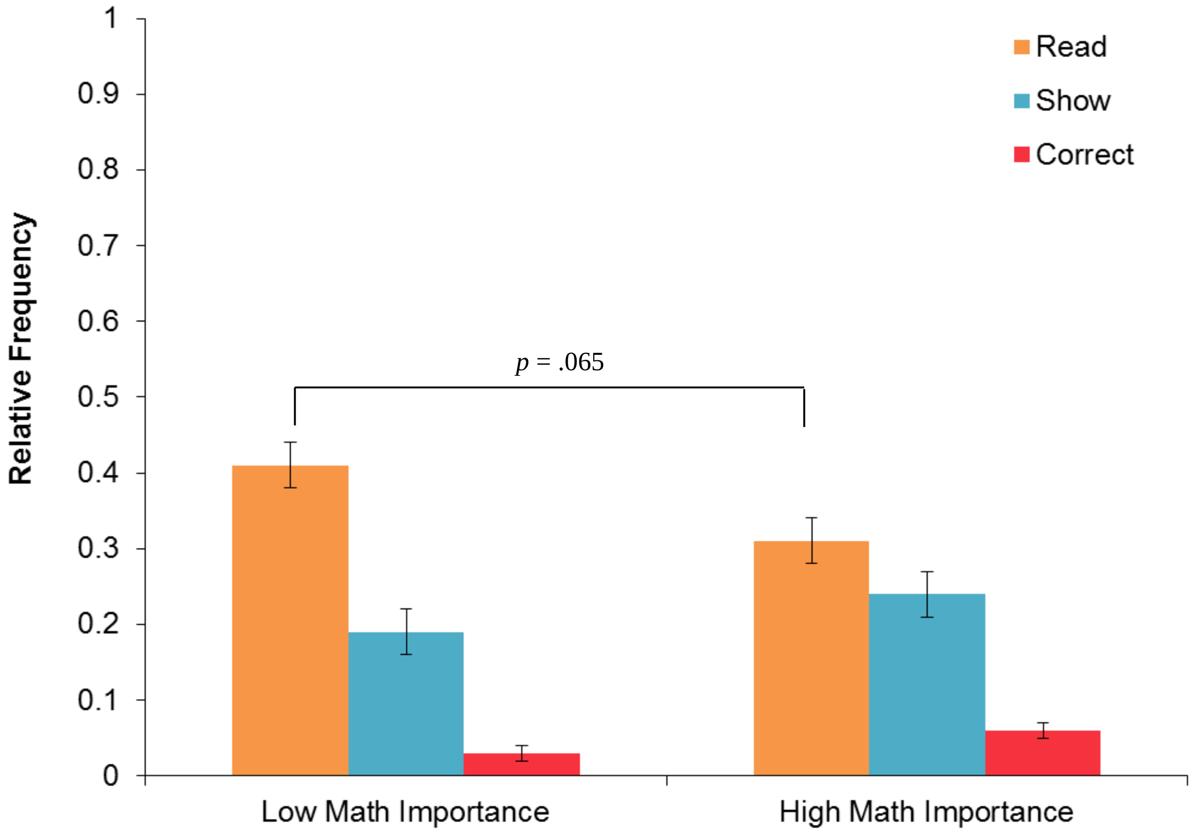



*Figure 8.* Mothers who valued math importance less used reading as a scaffolding strategy slightly more often than mothers who valued math importance more. $F(1.86,66.90) = 2.66$, $p = .082$, $\eta^2_p = .07$. Math importance was measured using the math item in the Perceptions of Importance subscale (Simpkins et al., 2011). Mothers' scores were divided into high and low groups using a median split (median = 6.00).

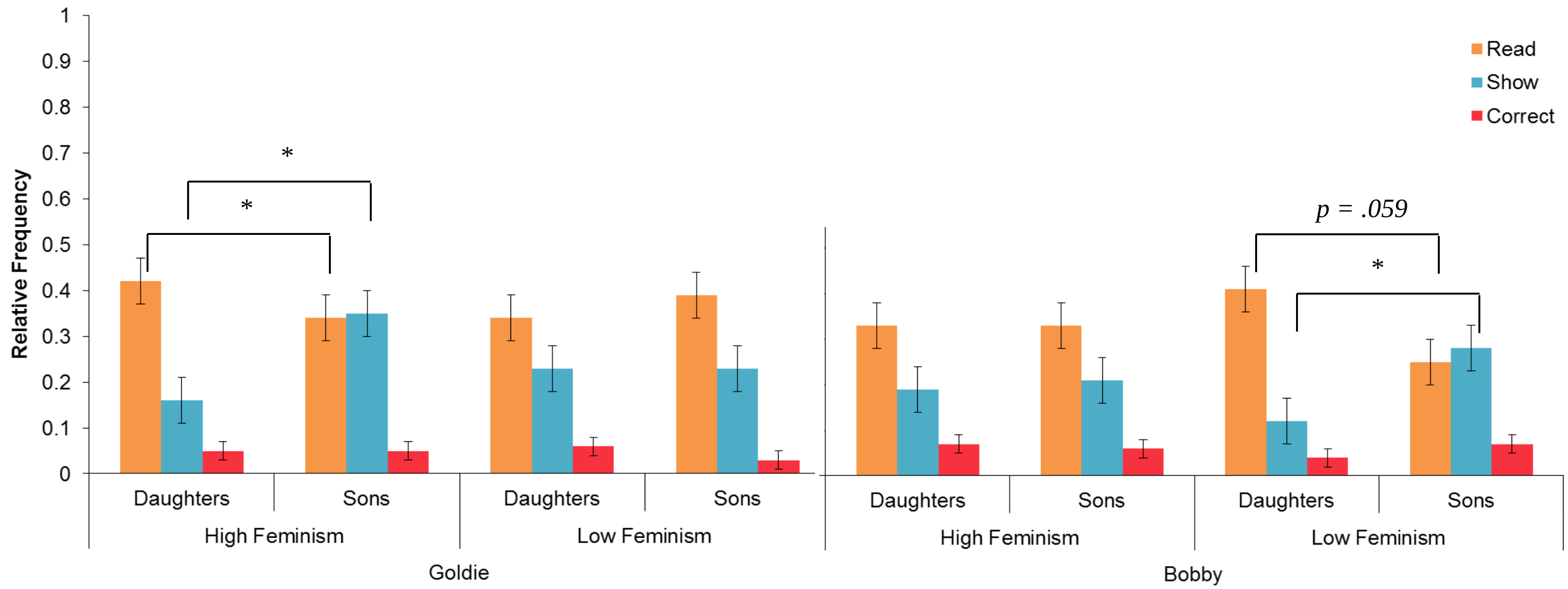



*Figure 9.* For dyads playing with *GoldieBlox*, highly feminist mothers demonstrated building more to sons than to daughters and read to daughters more than to sons (left). For dyads playing with *BobbyBlox*, low feminist mothers showed the same pattern, demonstrating building more often to sons than to daughters and reading to daughters more than to sons (right), $F(1.96,70.43) = 6.66$, $p = .002$, , $\eta^2_p = .16$. Feminism was measured using Becker and Wagner's (2009) Collective Action scale.  Mothers were divided into high and low feminism groups using a median split (median = 2.20).

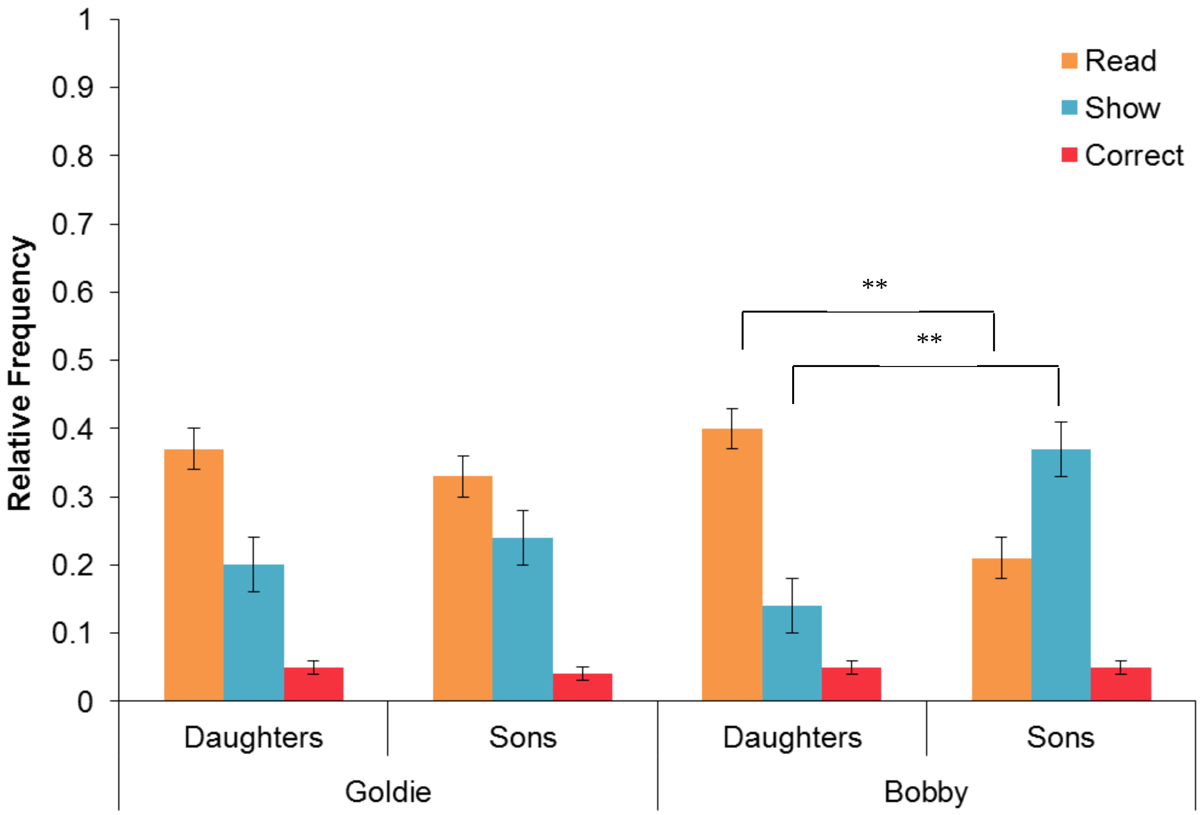



*Figure 10.* Mothers were especially gender-typed in their scaffolding strategy use with *BobbyBlox*, reading more to daughters than to sons and demonstrating building to sons more than to daughters, $F(1.44,76.03) = 4.94$, $p = .018$, $\eta^2_p = .09$.

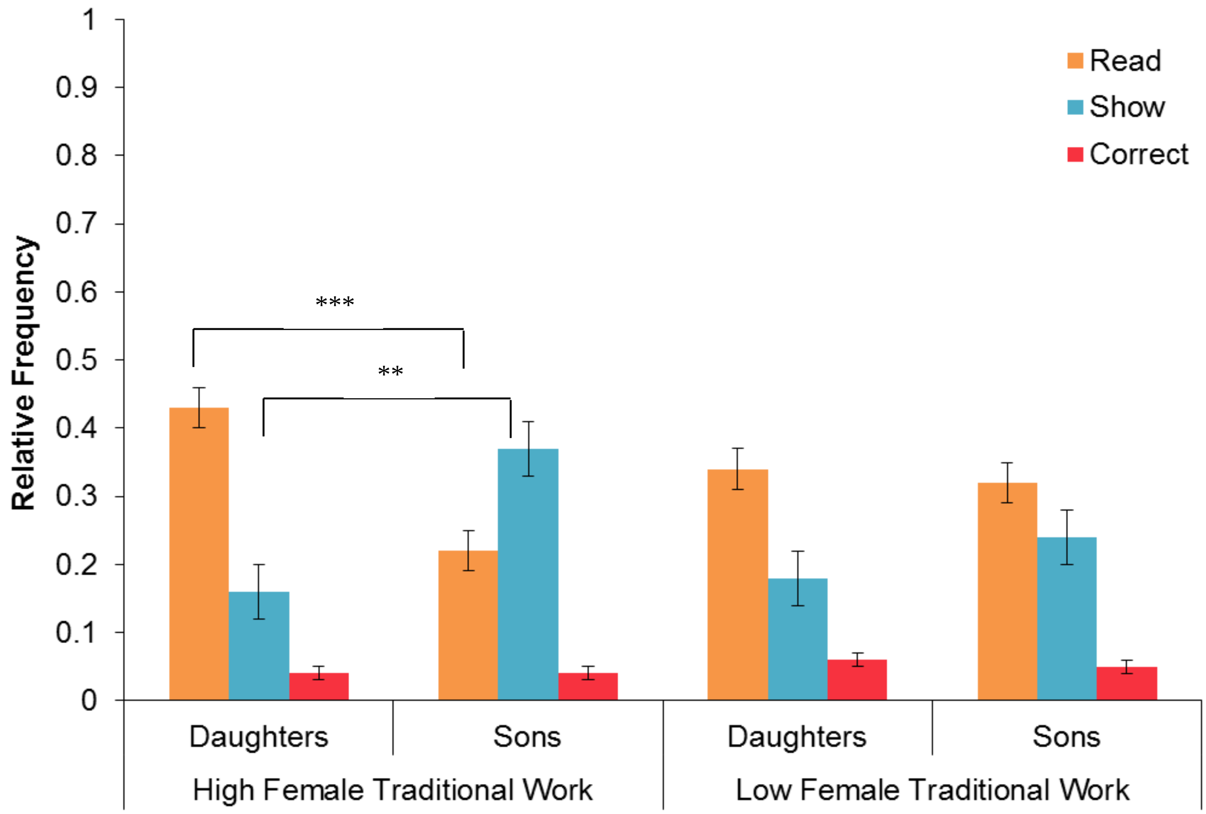



*Figure 11*. Mothers with highly traditionally feminine work were especially gender-typed in their use of scaffolding strategies, reading more to daughters than to sons and demonstrating building more to sons than to daughters, $F(1.44,76.03) = 4.94$, $p = .023$, $\eta^2_p = .08$. Mothers' work traditionality was the sum of three standardized scores: occupational femininity, occupational prestige, and hours employed. Composite scores were then divided in to high and low traditionality groups using a median split (median = 1.17).

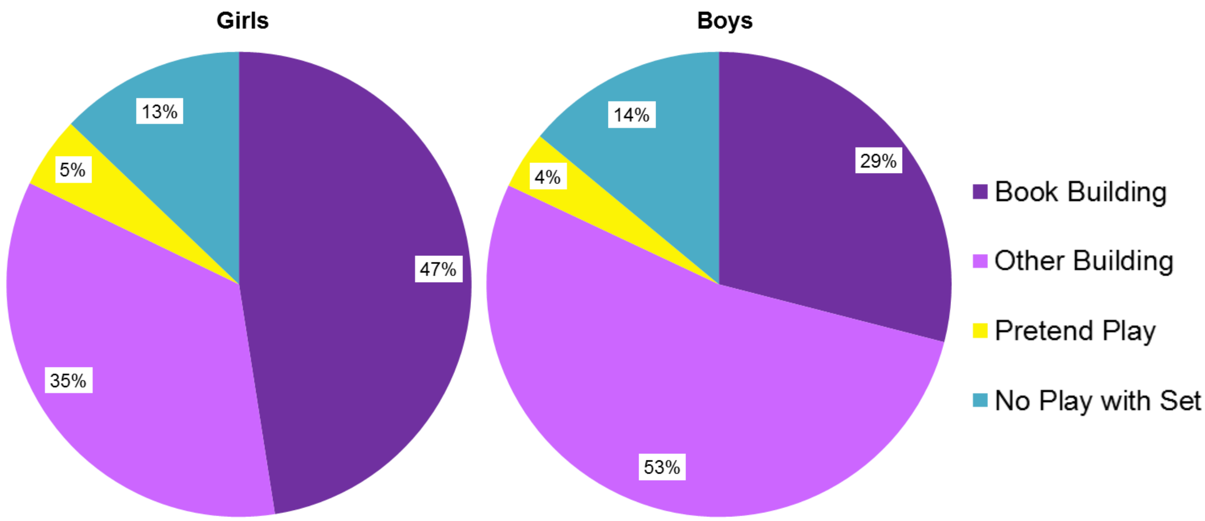



*Figure 12.* During free play, children engaged in both types of building more than any other activity, $F(1.78,98.12) = 22.47$, $p < .001$, $\eta^2_p = .29$. Girls engaged in more building guided by the book than did boys, whereas boys engaged in more building unrelated to the book than did girls, $F(1.78,98.12) = 3.39$, $p = .043$, $\eta^2_p = .06$.

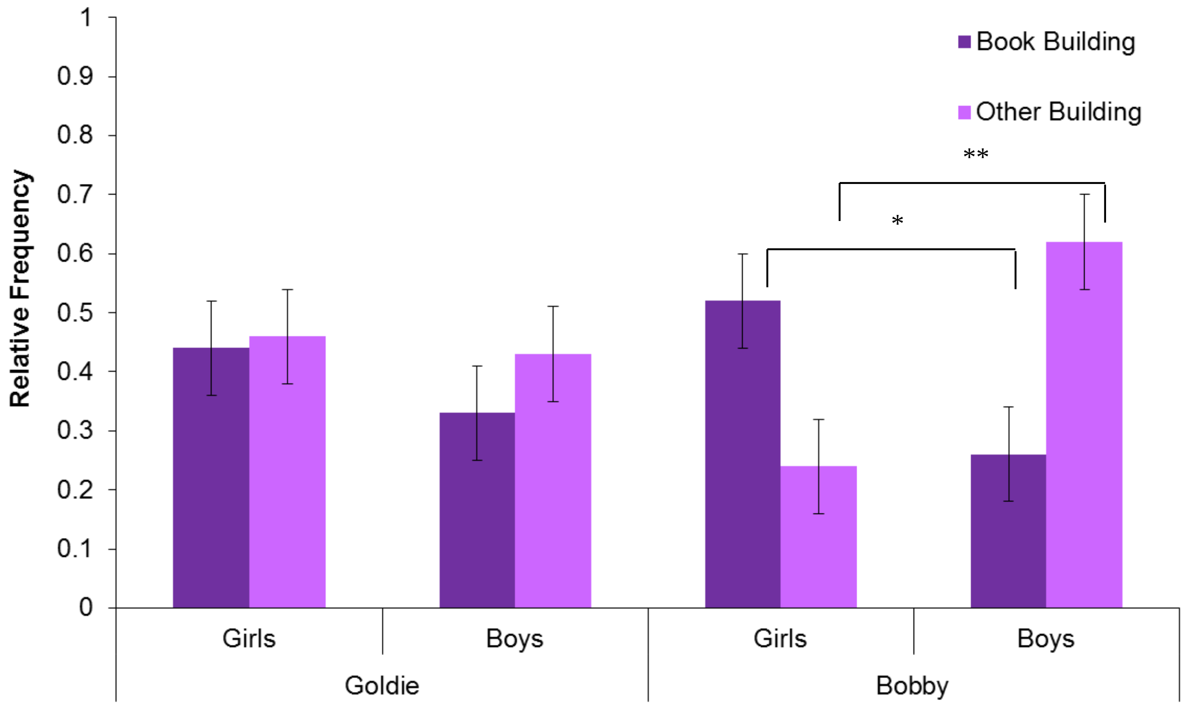



*Figure 13.* During free play, girls especially used the book to guide play with *BobbyBlox* more so than boys did. Boys built without the book while playing with *BobbyBlox* more so than girls did, $F(1.78,98.12) = 3.39$, $p = .048$, , $\eta^2_p = .05$.

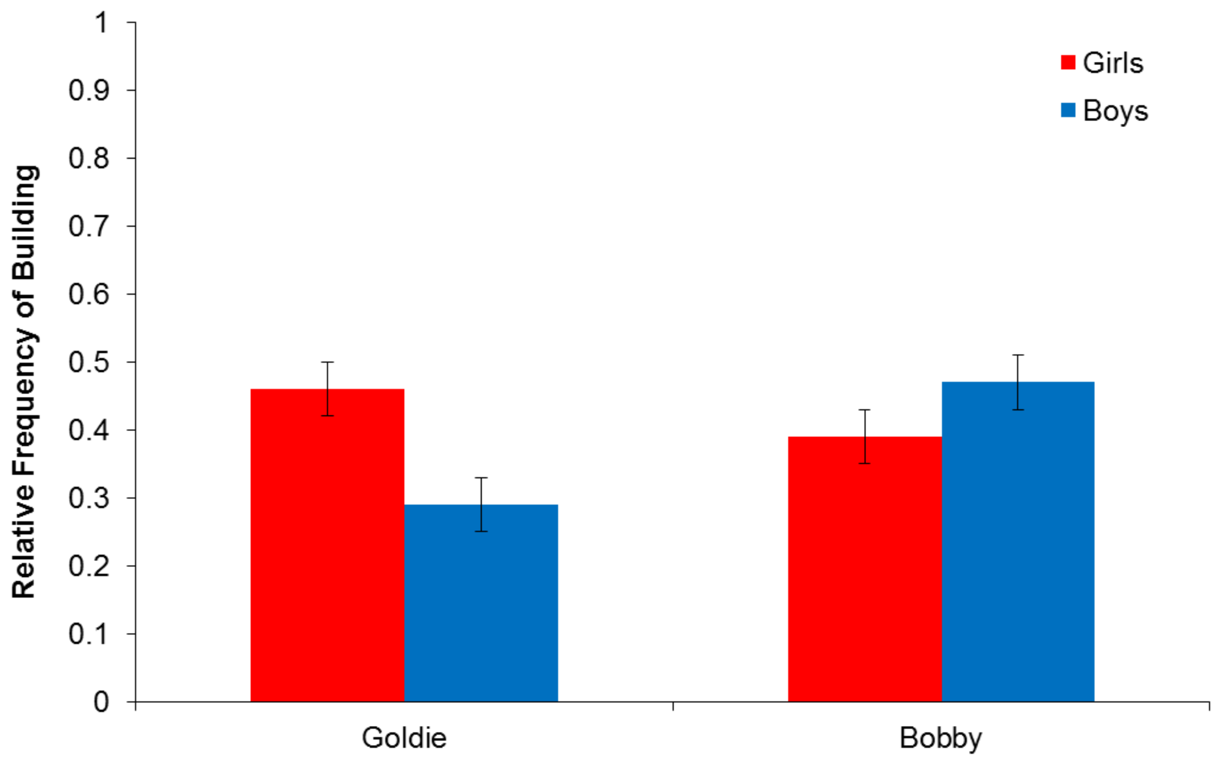



*Figure 14.* During free play, girls built more overall (both with and without the book) with *GoldieBlox* than with *BobbyBlox*. Boys built more overall with *BobbyBlox* than with *GoldieBlox*, $F(1,27) = 8.24$, $p = .008$, $\eta^2_p = .23$.

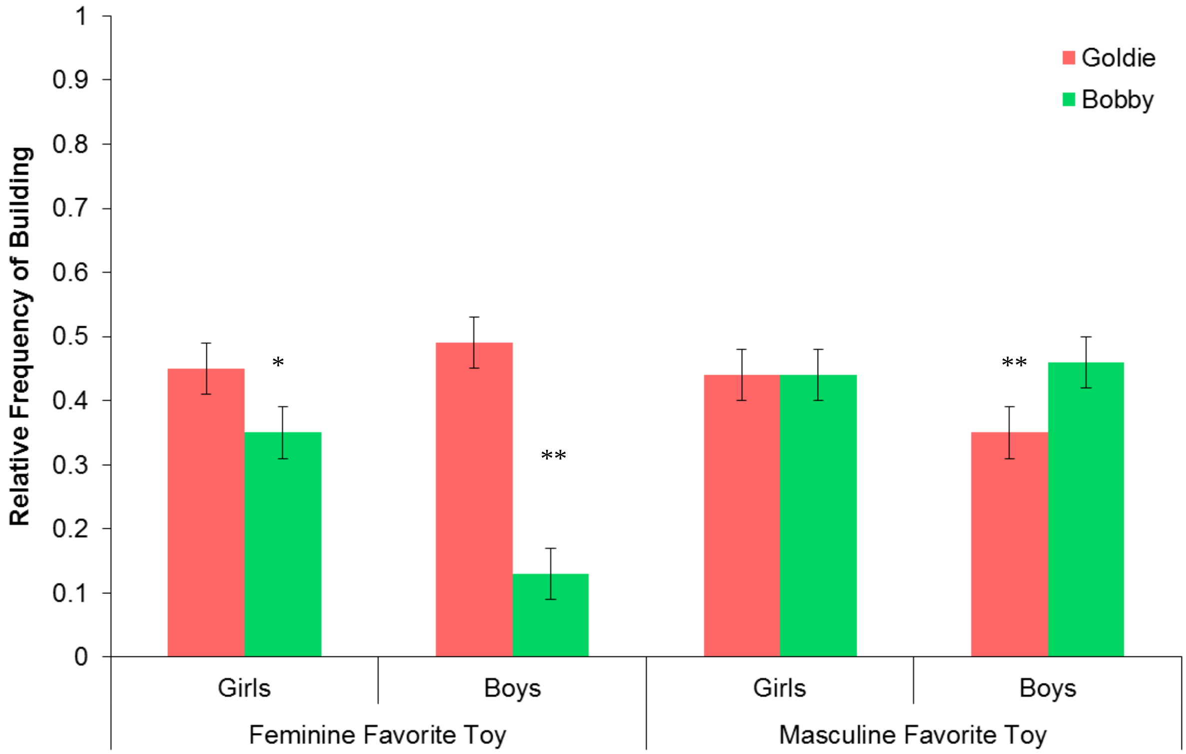



*Figure 15.* During free play, children with a feminine favorite toy built longer with *GoldieBlox* than did children with a masculine favorite toy. Children with a masculine favorite toy built longer with *BobbyBlox* than did children with a feminine favorite toy, $F(1,51) = 13.86$, $p < .001$, $\eta^2_p = .21$. Additionally, both girls and boys with a feminine favorite toy built more with *GoldieBlox* than *BobbyBlox* when they had a feminine favorite toy. Boys built more with *BobbyBlox* than *GoldieBlox* when they had a masculine favorite toy, $F(1,51) = 5.98$, $p = .018$, $\eta^2_p = .11$. Children reported their favorite toys during the laboratory session and they were coded for masculinity-femininity. Scores were then divided into masculine and feminine groups using a median split (median = 2.99).

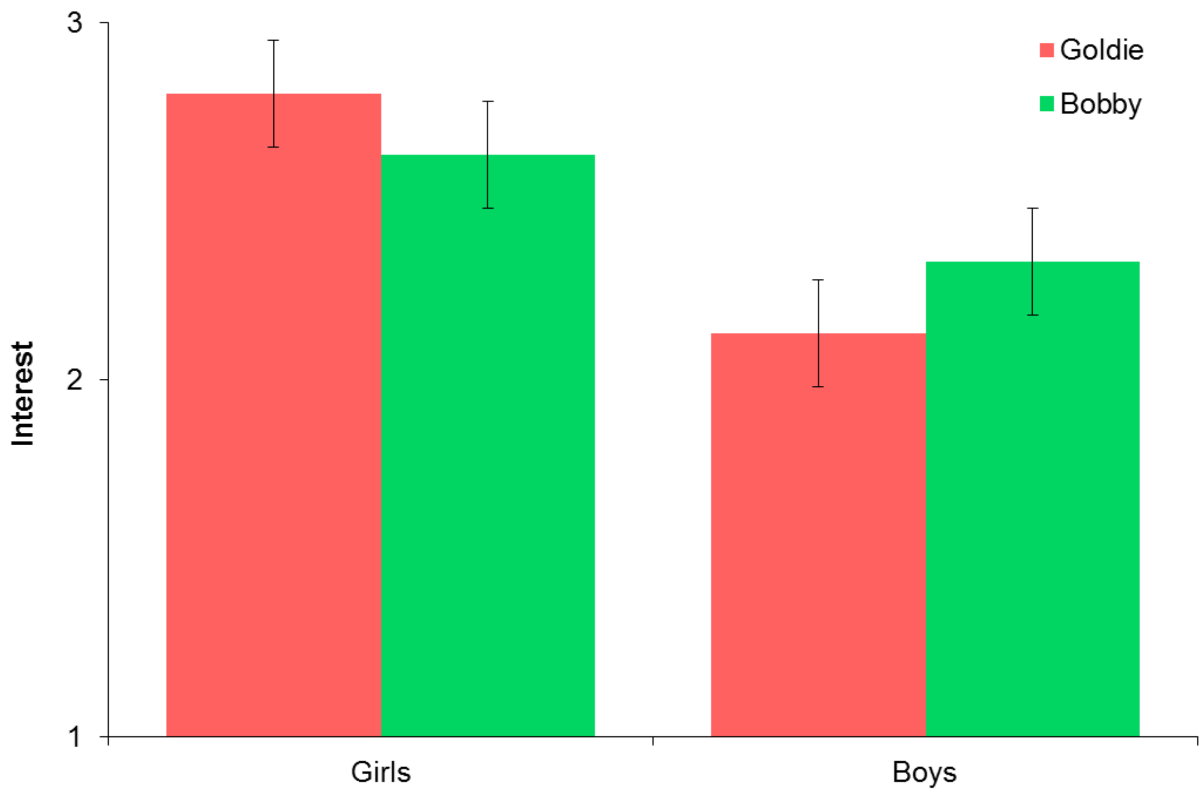



*Figure 16.* Girls were more interested in both toys following play, regardless of condition. Interest was assessed by asking children “How much did you enjoy playing with this toy today?” Response were made on a 3-point Likert scale of “Not at all” (1), "Some" (2), and “A lot” (3).

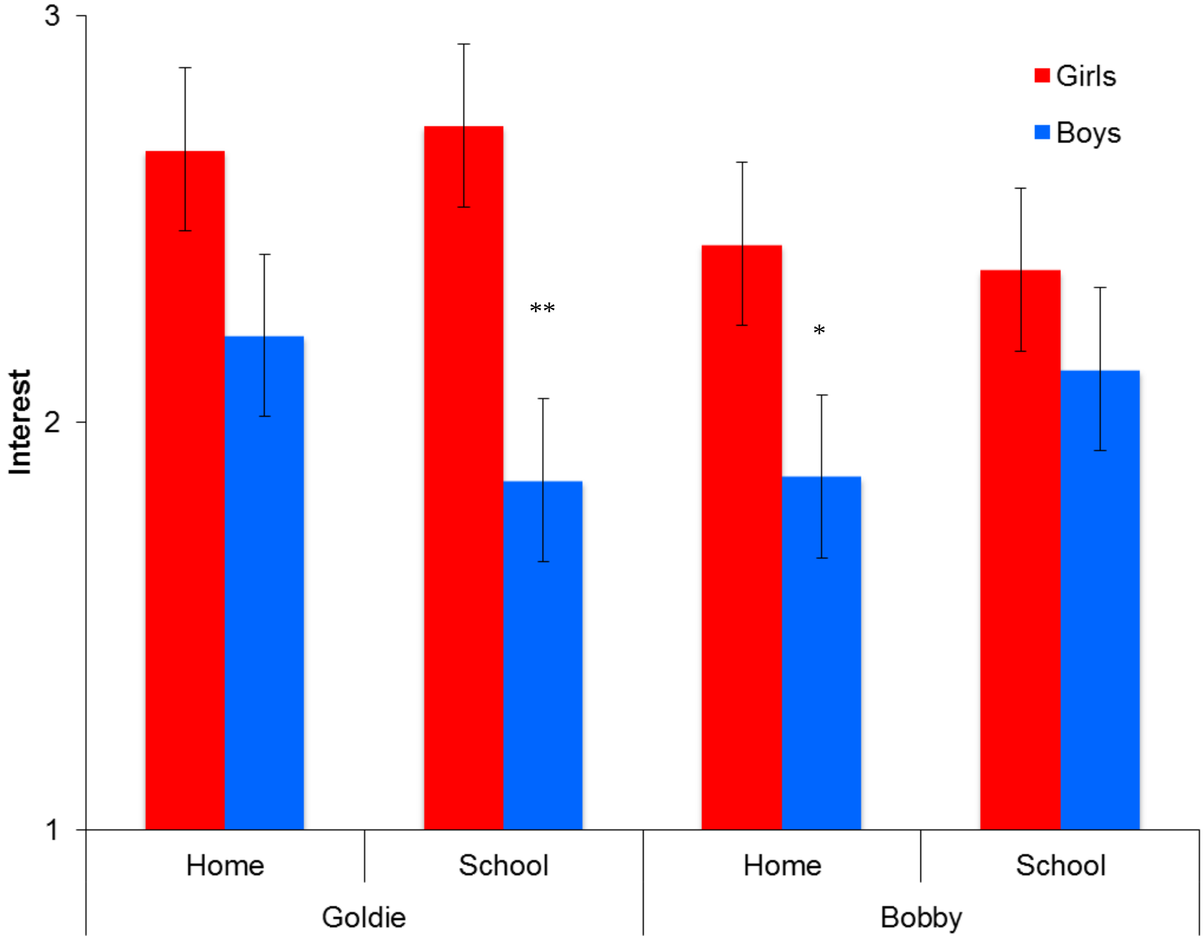



*Figure 17.* Boys were less interested in having *GoldieBlox* at school than were girls. Girls were more interested in having *BobbyBlox* at home than were boys, $F(1,56) = 3.30$, $p = .075$, $\eta^2_p = .06$. Interest was measured by asking children "How much would you like to play with this toy at home?" and "How much would you like to play with this toy at school?" Response were made on a 3-point Likert scale of "Not at all" (1), "Some" (2), and "A lot" (3).

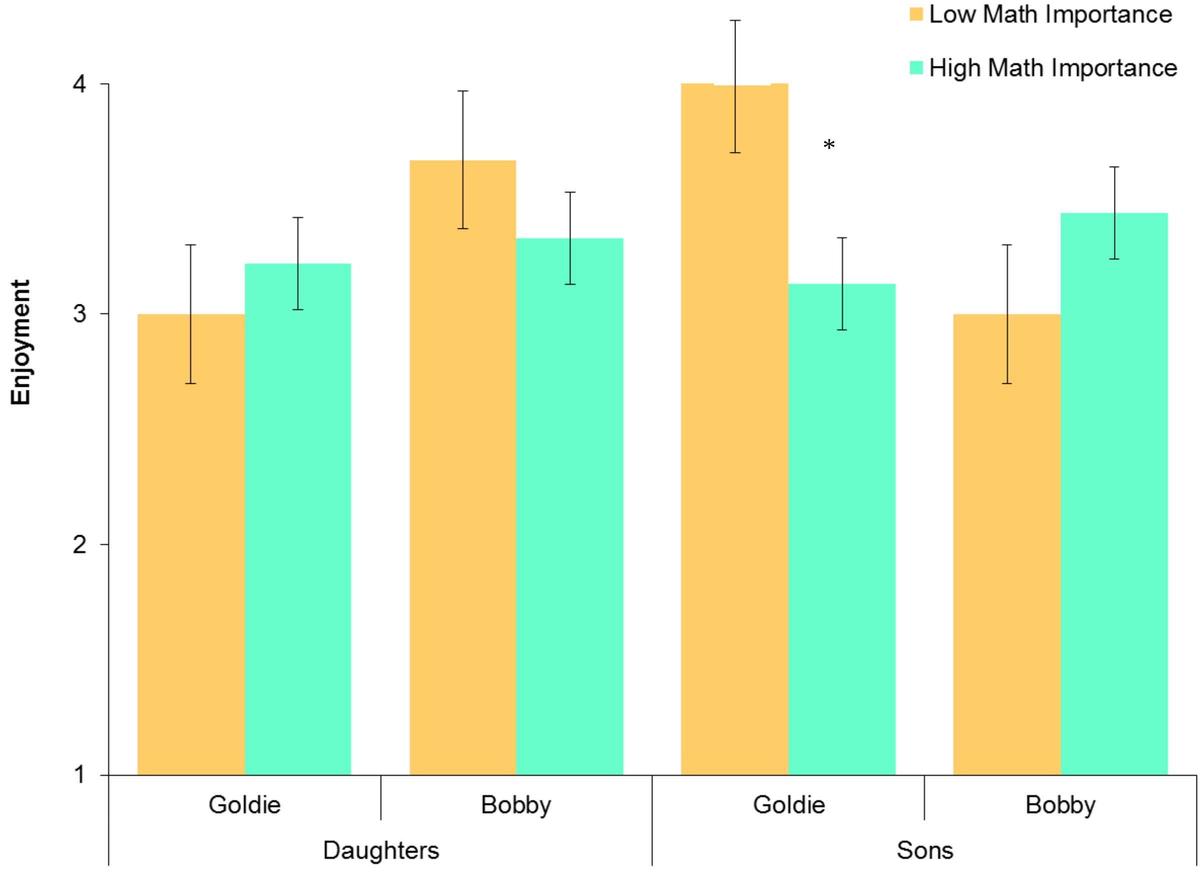



*Figure 18*. Mothers who valued math highly rated *GoldieBlox* as less enjoyable when playing with their son, as compared to mothers who did not value math as highly, $F(1,36) = 5.40$, $p = .026$, $\eta^2_p = .13$. Math importance was measured using the math item in the Perceptions of Importance subscale (Simpkins et al., 2011). Mothers' scores were divided into high and low groups using a median split (median = 6.00).

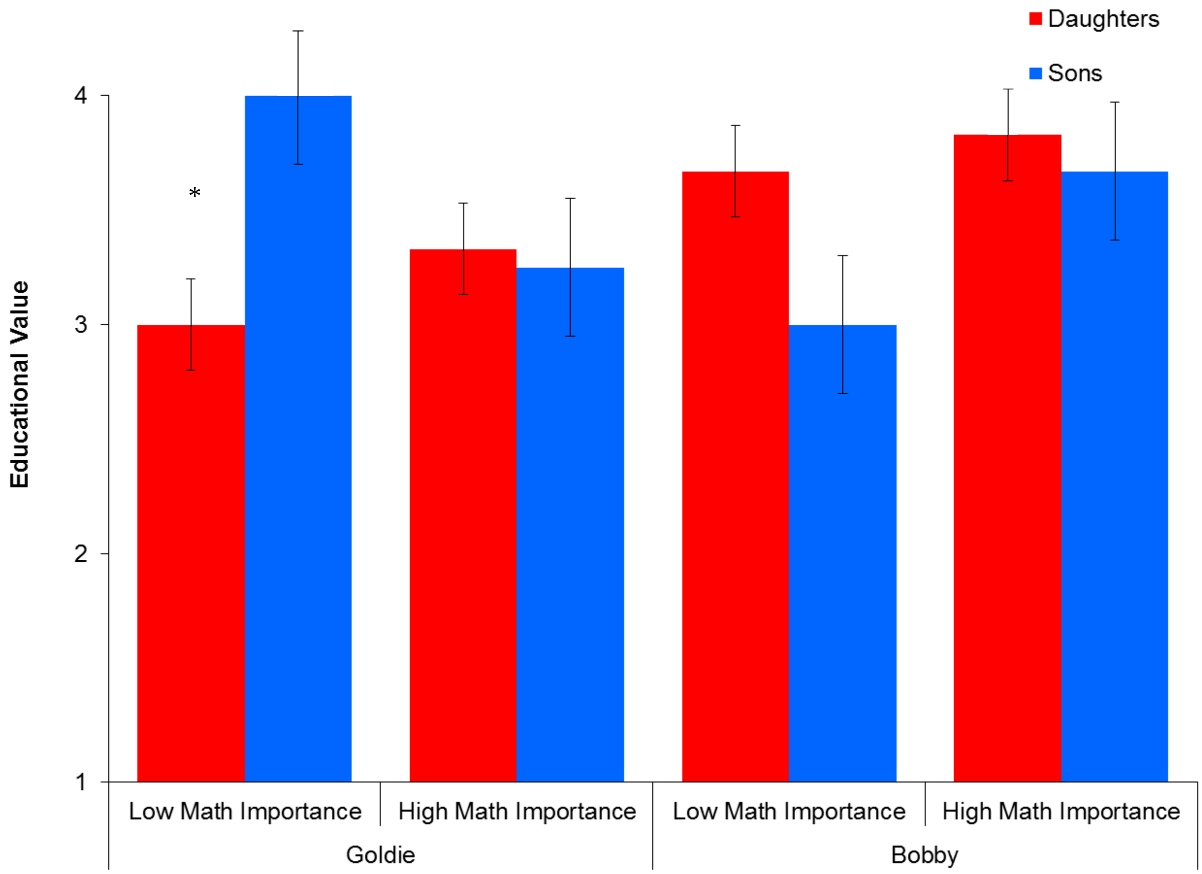



*Figure 19.* Mothers who valued math somewhat less rated *GoldieBlox* as more educational for sons than for daughters, $F(1,36) = 3.74$, $p = 0.62$, $\eta^2_p = .09$. Math importance was measured using the math item in the Perceptions of Importance subscale (Simpkins et al., 2011). Mothers' scores were divided into high and low groups using a median split (median = 6.00).

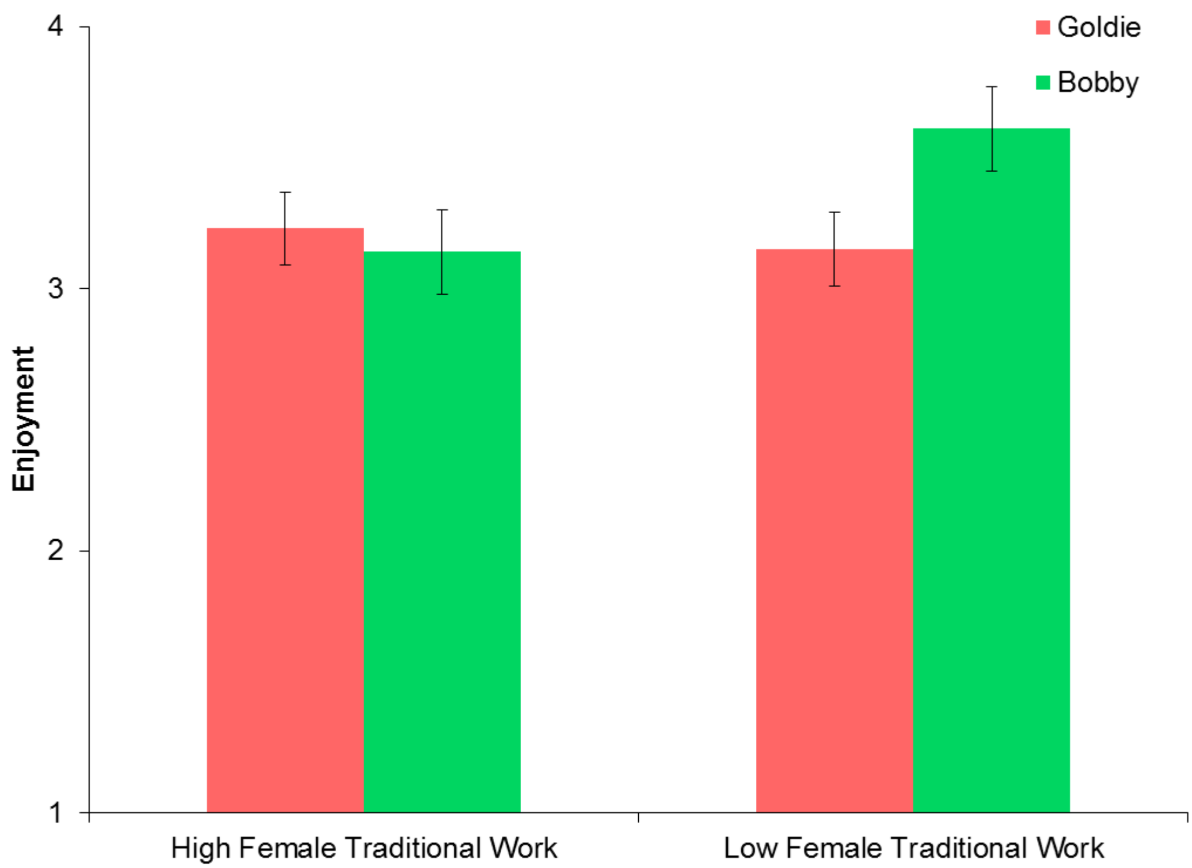



*Figure 20.* Mothers with less gender-traditional employment enjoyed *BobbyBlox* more than *GoldieBlox*, $F(1,53) = 3.07$, $p = .080$, $\eta^2_p = .06$. Mothers' work traditionality was the sum of three standardized scores: occupational femininity, occupational prestige, and hours employed. Composite scores were then divided in to high and low traditionality groups using a median split (median = 1.17).

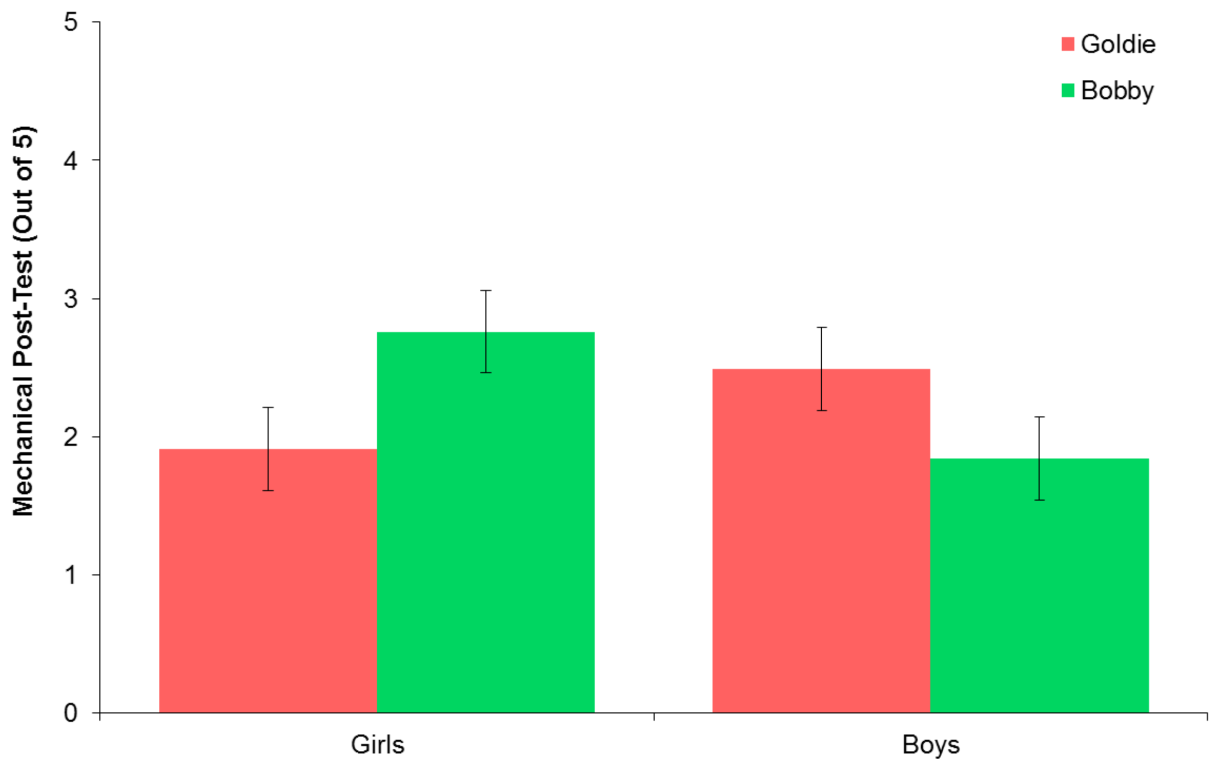



*Figure 21.* Girls learned more from *BobbyBlox* than from *GoldieBlox* whereas boys learned more from *GoldieBlox* than from *BobbyBlox*, $F(1,53) = 5.27$, $p = .025$, $\eta^2_p = .09$. Mechanical learning was measured by asking children to build a series of 5 belt-drive constructions in increasing complexity.

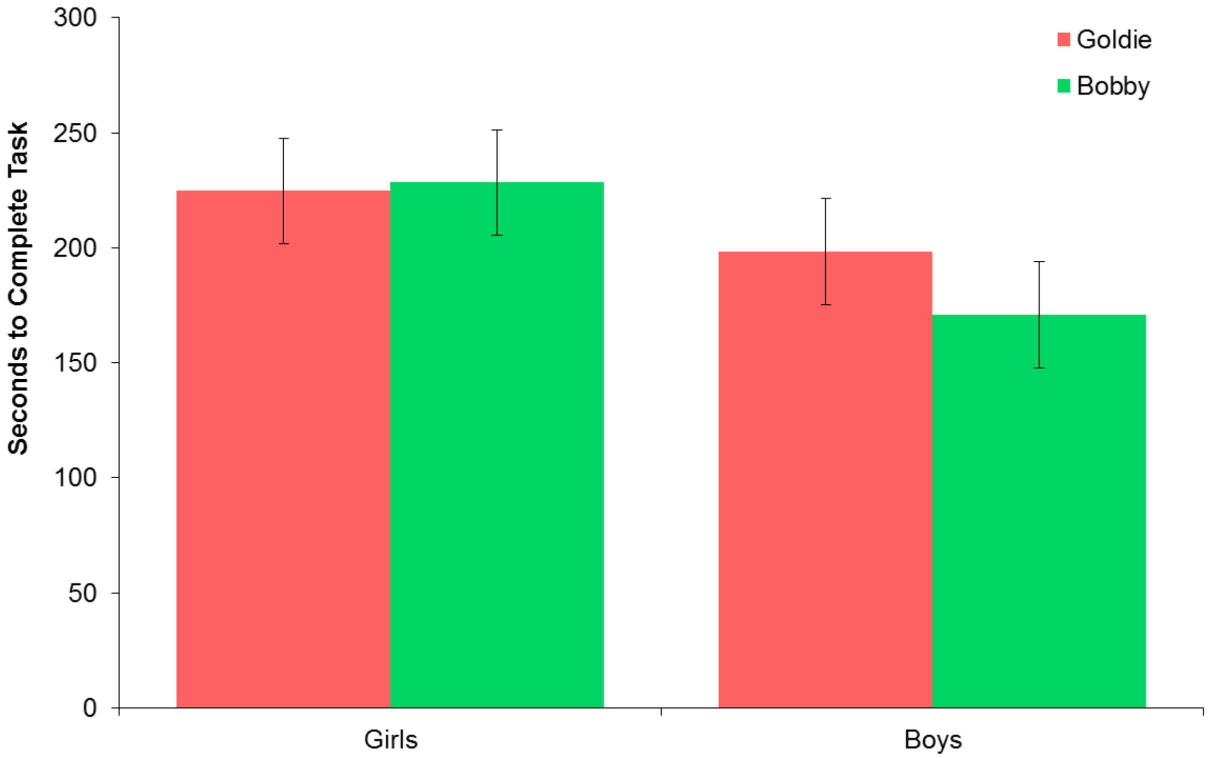



*Figure 22.* On average, boys completed the *Easy Transfer Task* faster than girls did, $F(1,56) = 4.15$, $p = .046$, $\eta^2_p = .07$. The task involved building a construction from a photograph (Appendix D) and the score was the time to completely build the structure.

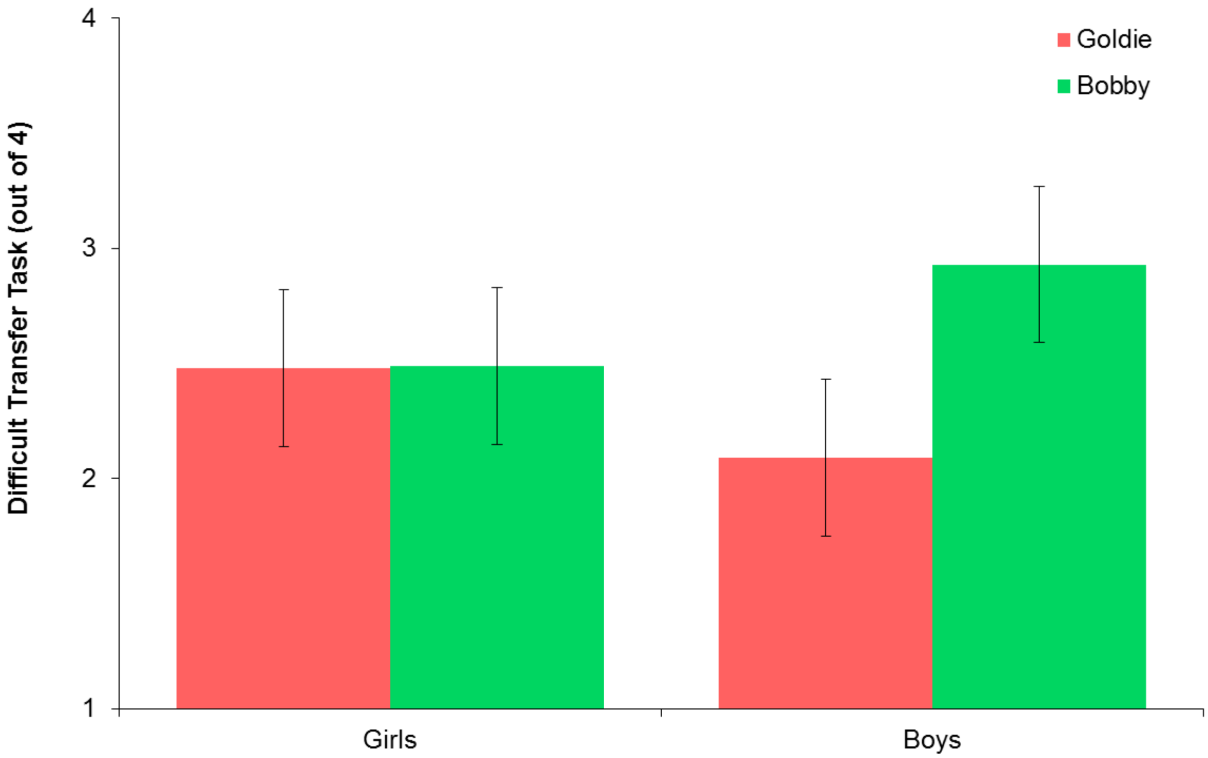



*Figure 23*. There were no effects of gender or condition on performance on the *Difficult Transfer Task* beyond the main effect of age, $F(1,54) = 14.69$, $p < .001$, $\eta^2_p = .21$. Older children outperformed younger children. This task involved building a difficult construction from an image (Appendix E) and was scored for completion and accuracy on a 4-point scale.

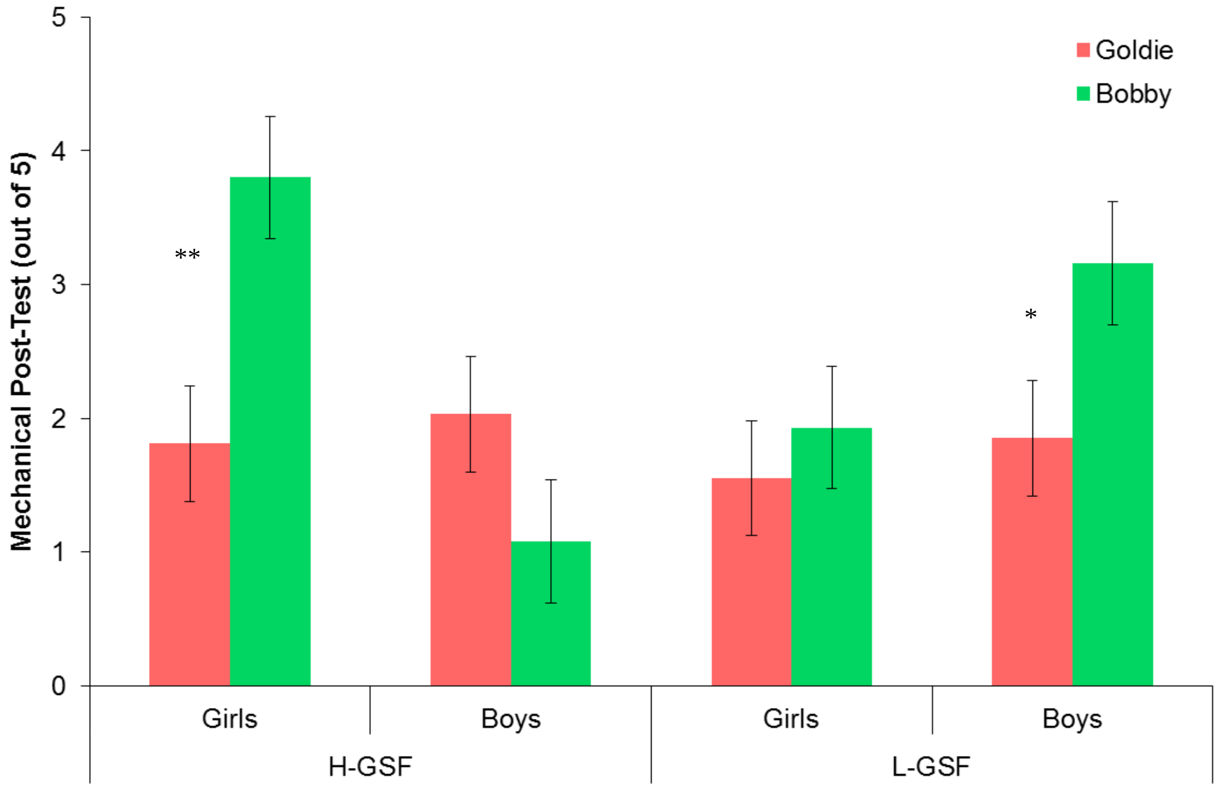



*Figure 24.* Gender salience interacted with gender and condition to affect mechanical learning. High salience (H-GSF) girls especially benefited from playing with *BobbyBlox* rather than *GoldieBlox*. Low salience (L-GSF) boys showed the same pattern, learning more from *BobbyBlox* than *GoldieBlox*, $F(1,27) = 7.61$, $p = .010$, $\eta^2_p = .22$. GSF level was determined by a median split on the GSF score (sum of three standardized scores: gendered affiliation, gender memory, and girl typicality; median = -0.14). Mechanical learning was measured by asking children to build a series of 5 belt-drive constructions in increasing complexity.

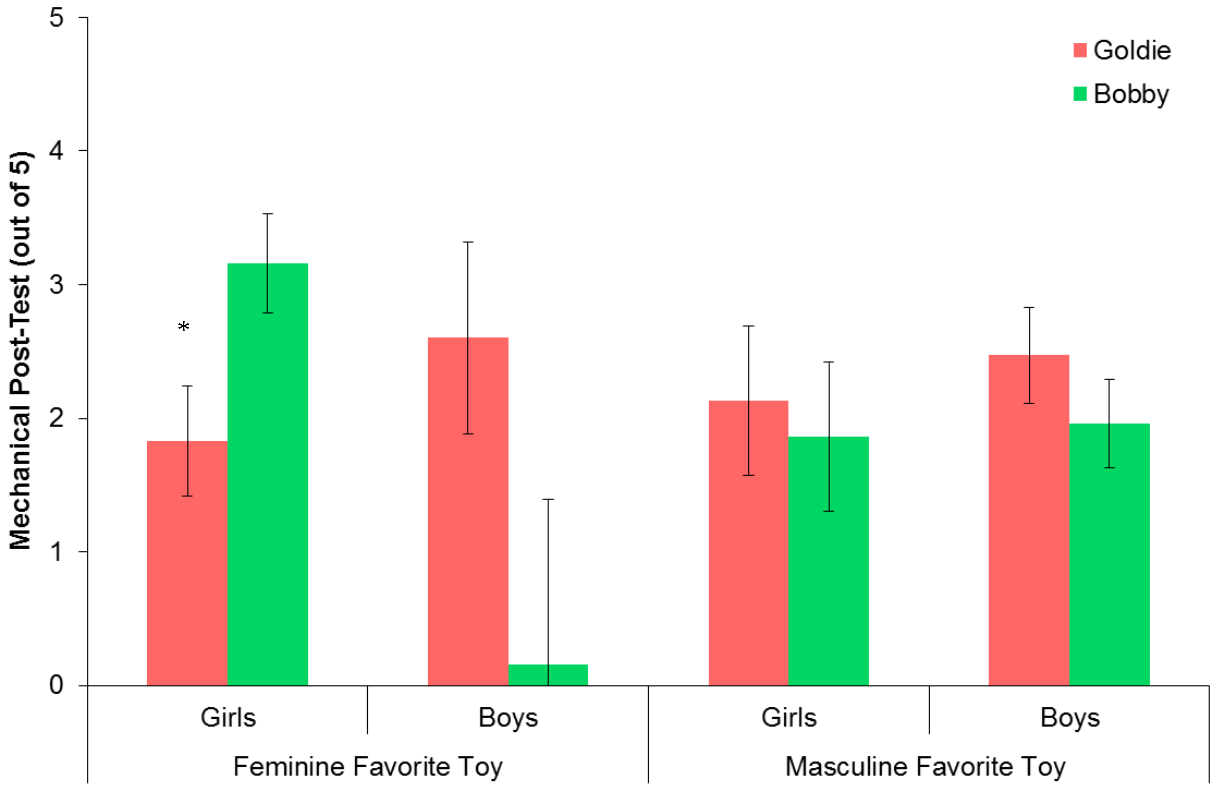



*Figure 25.* Girls with a feminine favorite toy especially benefited from playing with *BobbyBlox* rather than *GoldieBlox* in terms of mechanical learning outcomes. $F(1,52) = 3.91$, $p = .053$, $\eta^2_p = .07$. Children reported their favorite toys during the laboratory session and they were coded for masculinity-femininity. Scores were then divided into masculine and feminine groups using a median split (median = 2.99). Mechanical learning was measured by asking children to build a series of 5 belt-drive constructions in increasing complexity.

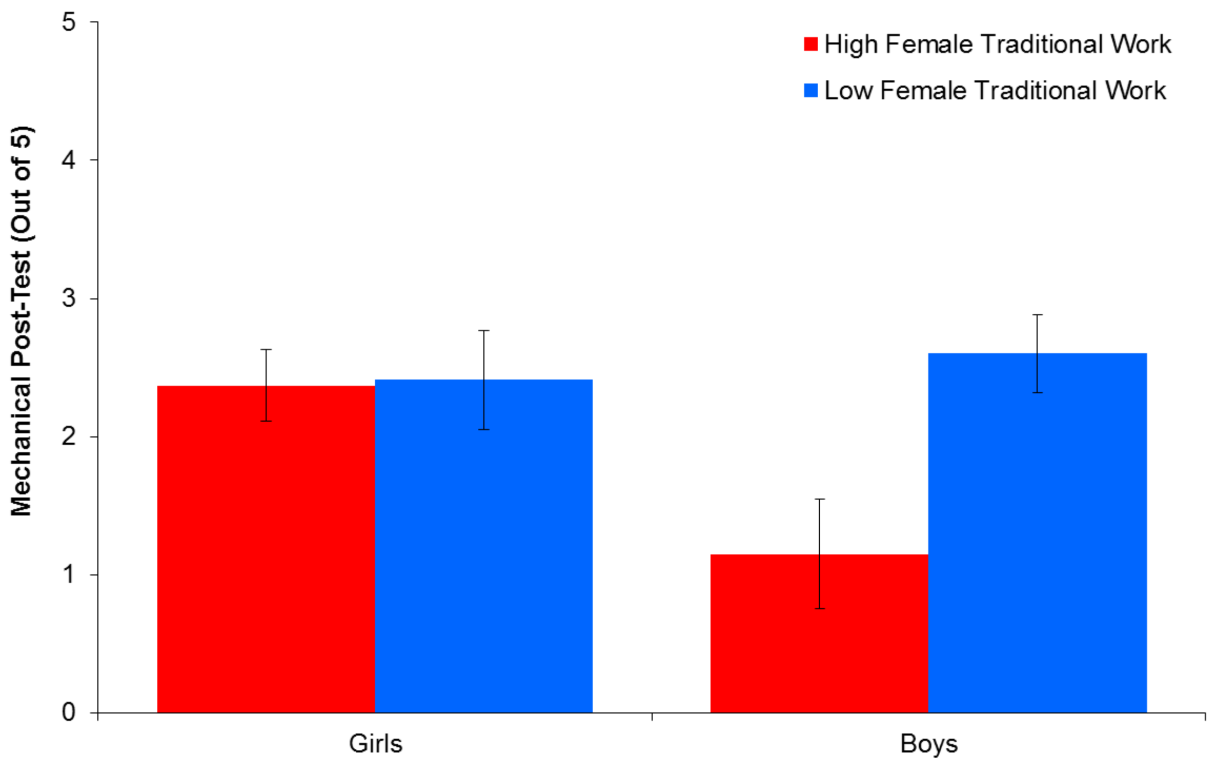



*Figure 26.* Children whose mothers held less gender-traditional employment learned more overall than children whose mothers held more-traditional employment, $F(1,52) = 7.43$, $p = .024$, $\eta^2_p = .09$. Boys whose mothers held less gender-traditional employment learned more than boys whose mothers held more gender-traditional employment, $F(1,52) = 4.80$, $p = .033$, $\eta^2_p = .08$. Mothers' work traditionality was the sum of three standardized scores: occupational femininity, occupational prestige, and hours employed. Composite scores were then divided in to high and low traditionality groups using a median split (median = 1.17). Mechanical learning was measured by asking children to build a series of 5 belt-drive constructions in increasing complexity.

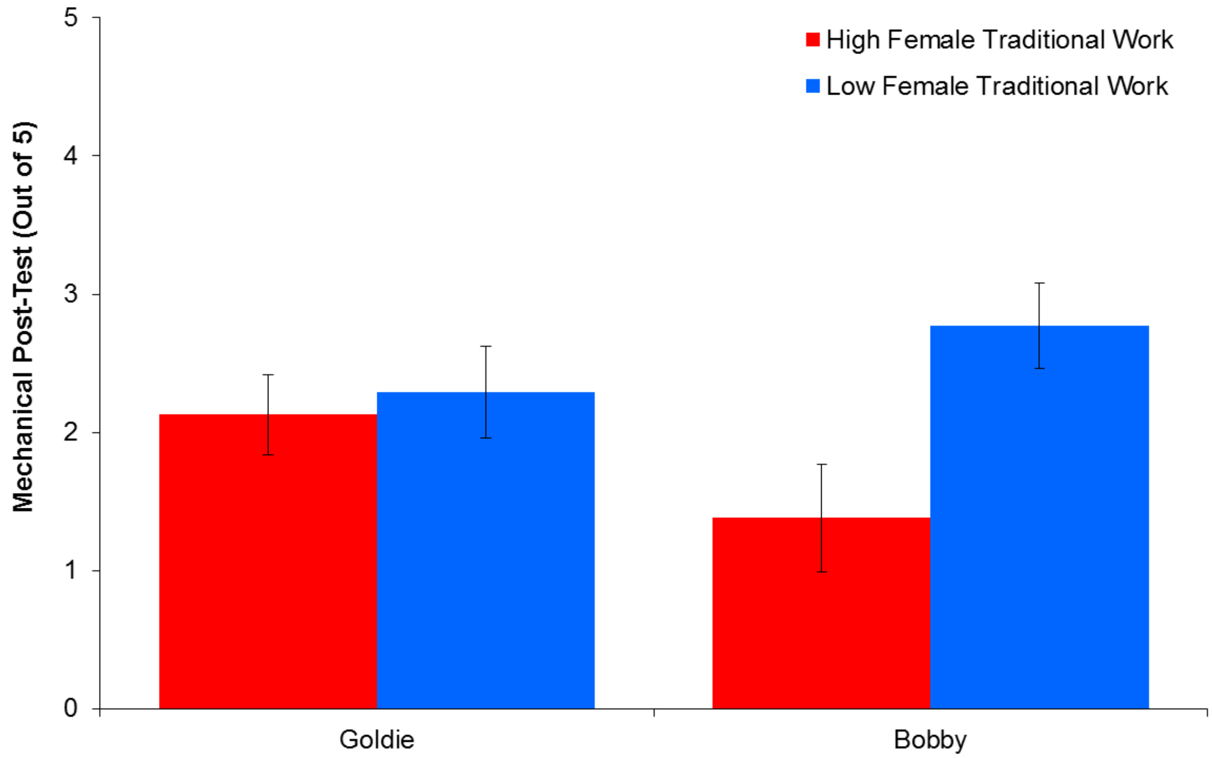



*Figure 27.* Children whose mothers held less gender-traditional employment learned more overall than children whose mothers held more-traditional employment, $F(1,52) = 7.43$, $p = .024$, $\eta^2_p = .09$. Among dyads playing with *BobbyBlox*, children whose mothers held less gender-traditional employment learned more than children whose mothers held more gender-traditional employment, $F(1,52) = 3.35$, $p = .073$, $\eta^2_p = .06$. Mothers' work traditionality was the sum of three standardized scores: occupational femininity, occupational prestige, and hours employed. Composite scores were then divided in to high and low traditionality groups using a median split (median = 1.17). Mechanical learning was measured by asking children to build a series of 5 belt-drive constructions in increasing complexity.

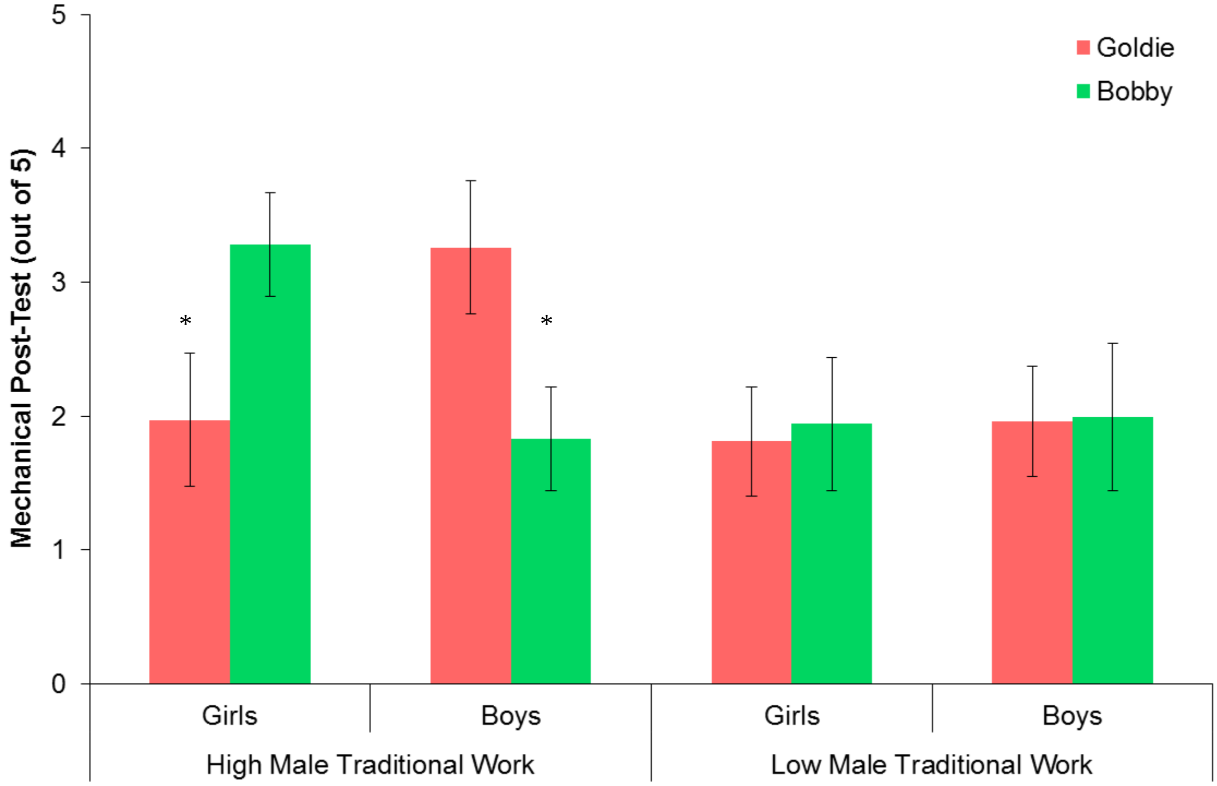



*Figure 28.* Children whose fathers held more gender-traditional employment learned more overall than did children whose fathers held less-traditional employment, $F(1,52) = 3.90$, $p = .054$, $\eta^2_p = .07$. Girls whose fathers had gender-traditional jobs benefitted especially from playing with *BobbyBlox* versus *GoldieBlox*, while boys whose fathers had gender-traditional jobs benefitted especially from playing with *GoldieBlox*, $F(1,52) = 4.06$, $p = .049$, $\eta^2_p = .07$. Fathers' work traditionality was the sum of three standardized scores: occupational masculinity, occupational prestige, and hours employed. Composite scores were then divided in to high and low traditionality groups using a median split (median = 0.17). Mechanical learning was measured by asking children to build a series of 5 belt-drive constructions in increasing complexity.

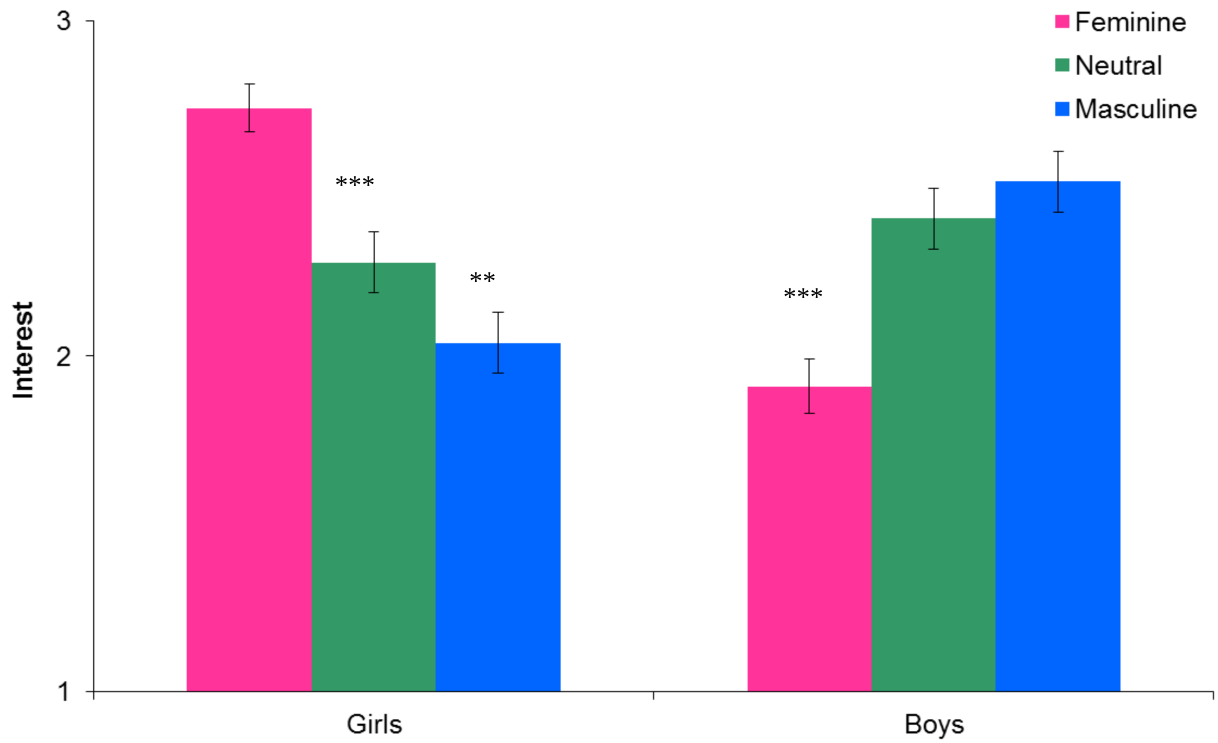



*Figure 29.* Children showed sex-typed interest in potential future play with other STEM toys. Girls preferred feminine toys to neutral toys, and neutral toys to masculine toys. Boys preferred neutral or masculine toys to feminine toys. Children rated the STEM toys depicted in Appendix F using a 3-point Likert scale of “Not at all” (1), "Some" (2), and “A lot” (3).

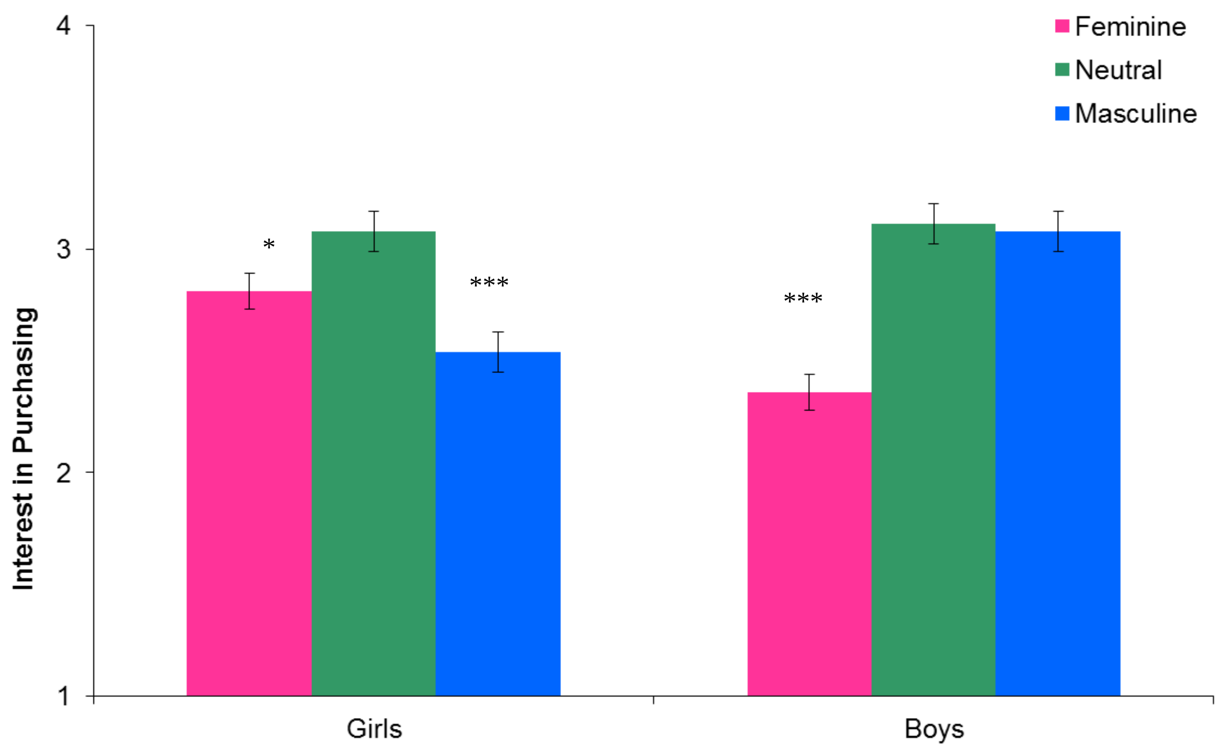



*Figure 30.* Mothers were most interested in purchasing neutral STEM toys overall, $F(2,114) = 24.76$, $p < .001$, $\eta^2_p = .30$. There was also a toy type by gender interaction, $F(2,114) = 22.89$, $p < .001$, $\eta^2_p = .29$. For daughters, mothers were more interested in neutral toys, then feminine toys, then masculine toys. For sons, mothers were equally interested in neutral and masculine toys, and least interested in feminine toys. Mothers rated the STEM toys depicted in Appendix F using a 4-point Likert scale from “Not at all interested” to “Very interested”.

# APPENDIX B

Experimental Script

Welcome to the Cognitive and Social Development lab! My name is Emily. Today you will get to try a new educational toy. We're going to do all of our play in this room and a couple activities in that room.

*Indicate 502 with door open*

Before we get started, I have a consent form for you to sign on behalf of you and your child. I will be videotaping this session so that we can look at it later, but all videos will be kept anonymous.

*Sign consent*

Now, I'm going to give mom/dad a chance to see the toy we will be playing with before you play together. You can go ahead and familiarize yourself. [Child's name] and I will talk about his/her favorite toys in the next room.

*Close door behind parent.*

*START CAMERA.*

*Seat child at computer, and administer first portion of "Lab survey – Child Survey." Administer child favorite toys, WLT.*

*Bring child back into room with parent.*

Now you two will get to play with this toy together. I will be in the other room, and I will come back in about 15 minutes.

*Set timer for 15 minutes.*

*Adjust camera if necessary. After 15 minutes, return to room.*

Great! Now I'm going to take [mom/dad] to the other room to answer some questions while you [child's name] play a little more by yourself. We will be right on the other side of the door.

*Start timer for 7 minutes.*

*Set up parent with survey "Lab session parent survey." After 5-7 minutes (or when child will no longer play), return to observation room but leave parent with survey, explain WLT.*

Now that you have had some time to play with this toy, I am going to ask you some questions about it. I will take some notes while you play so I can remember for later.

*Administer mechanical test.*

OK, I have one more toy for us to play with. I am going to show you a picture, and you get to try to build that. If you finish the first one, I have a second picture to use too. Look at this picture. Try to build exactly what you see in the picture.

*Take out Tinker Toys, administer easy transfer test and then difficult transfer test.*

We are going to do just a little more on the computer. I want to know what you thought of the toy we played with today, and how much you like some other toys.

*Take child back to computer. Return to web survey.*

Thank you both for coming in today! You will receive an emailed link in 5-7 days to the final online survey. Were you able to provide an email address in the web survey today?

(<u>ALTERNATIVELY</u>: This is the email I have on file for you. Is that a good address to send the survey to?**)**

*Note email address for survey, notice.*

Do you know anyone who might be interested in doing this study as well? We are looking for children 4-6 years old and one parent. Can I send you an email to forward to your friends who might be interested?

*Note email address for survey, notice. Give flyer or note contact information.*

I have a certificate for an ice cream cone for [child's name]! You are also both entered in the raffle to win a gift card to Toys R Us or one of the games from this study if you enjoyed it. If you both do the online survey, you will get an additional raffle entry for each part you complete, and a final bonus entry for completing everything! Do you need reimbursement for parking today?

*Reimburse. Give ice cream certificate.*

Thank you both for your help!

*<u>STOP CAMERA.</u>*

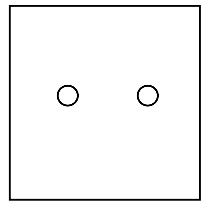
# APPENDIX C

Mechanical Learning Measure

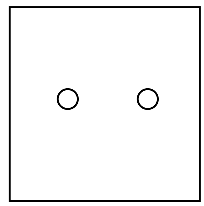
For this activity, I need you to build two wheels like you did before. Wind up the ribbon on one wheel. Now choose a character to put on each wheel.

**Character 1:** Nacho Benjamin Katinka Flavio Phil
**Character 2:** Nacho Benjamin Katinka Flavio Phil

How do you make (1) and (2) turn in the same direction?
Correct Incorrect Did not attempt

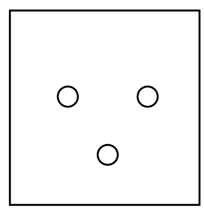


How do you make them turn in opposite directions?
Correct Incorrect Did not attempt

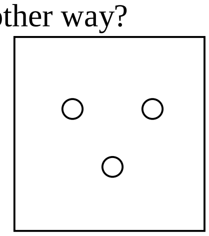


Now add a third wheel and pick another character.
**Character 3:** Nacho Benjamin Katinka Flavio Phil

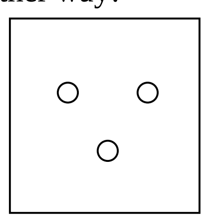
How do you make them all spin the same direction?
Correct 2/3 Incorrect Did not attempt



How do you make (1) and (2) turn together and (3) turn the other way?
Correct 2/3 Incorrect Did not attempt



How do you make (2) and (3) turn together and (1) turn the other way?
Correct 2/3 Incorrect Did not attempt

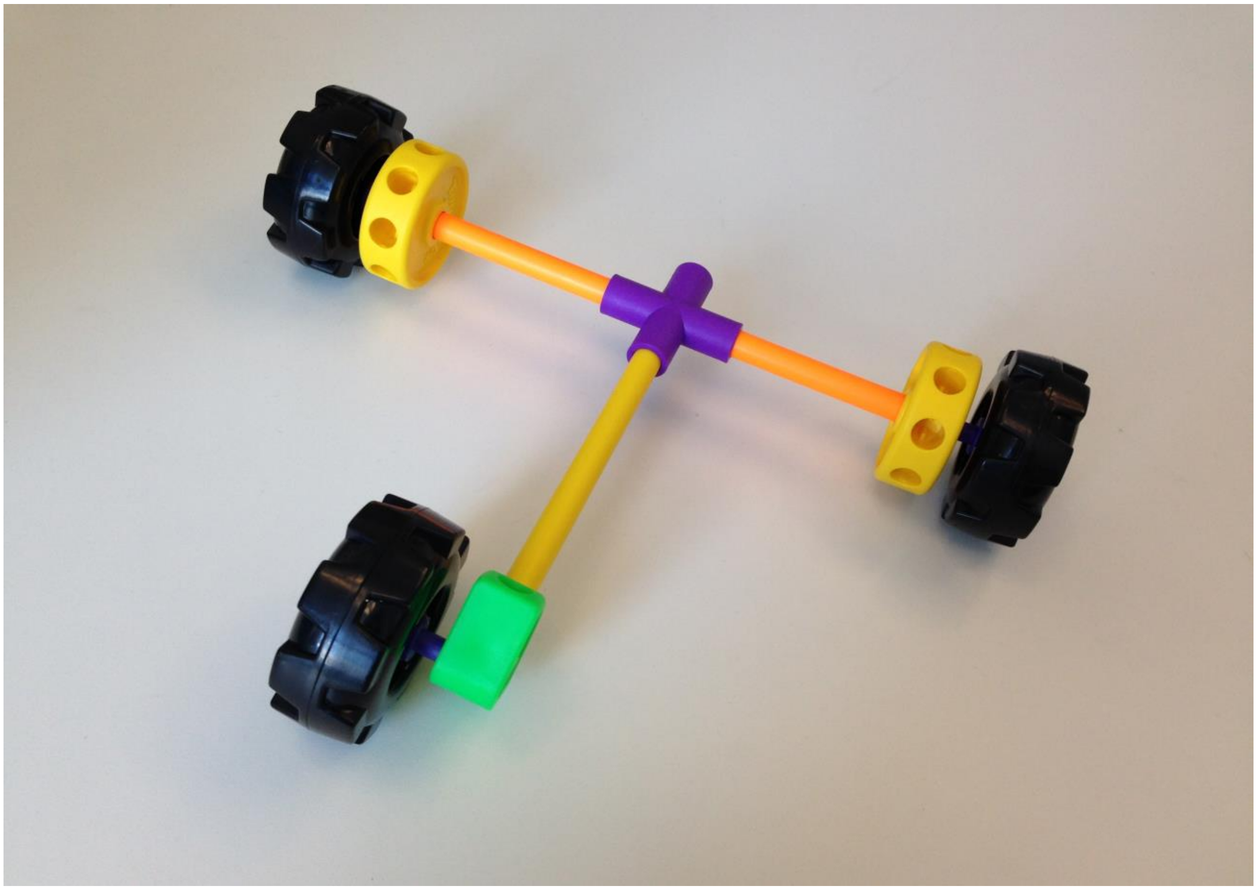


**Score out of 5 ____**

# APPENDIX D

Easy Transfer Task

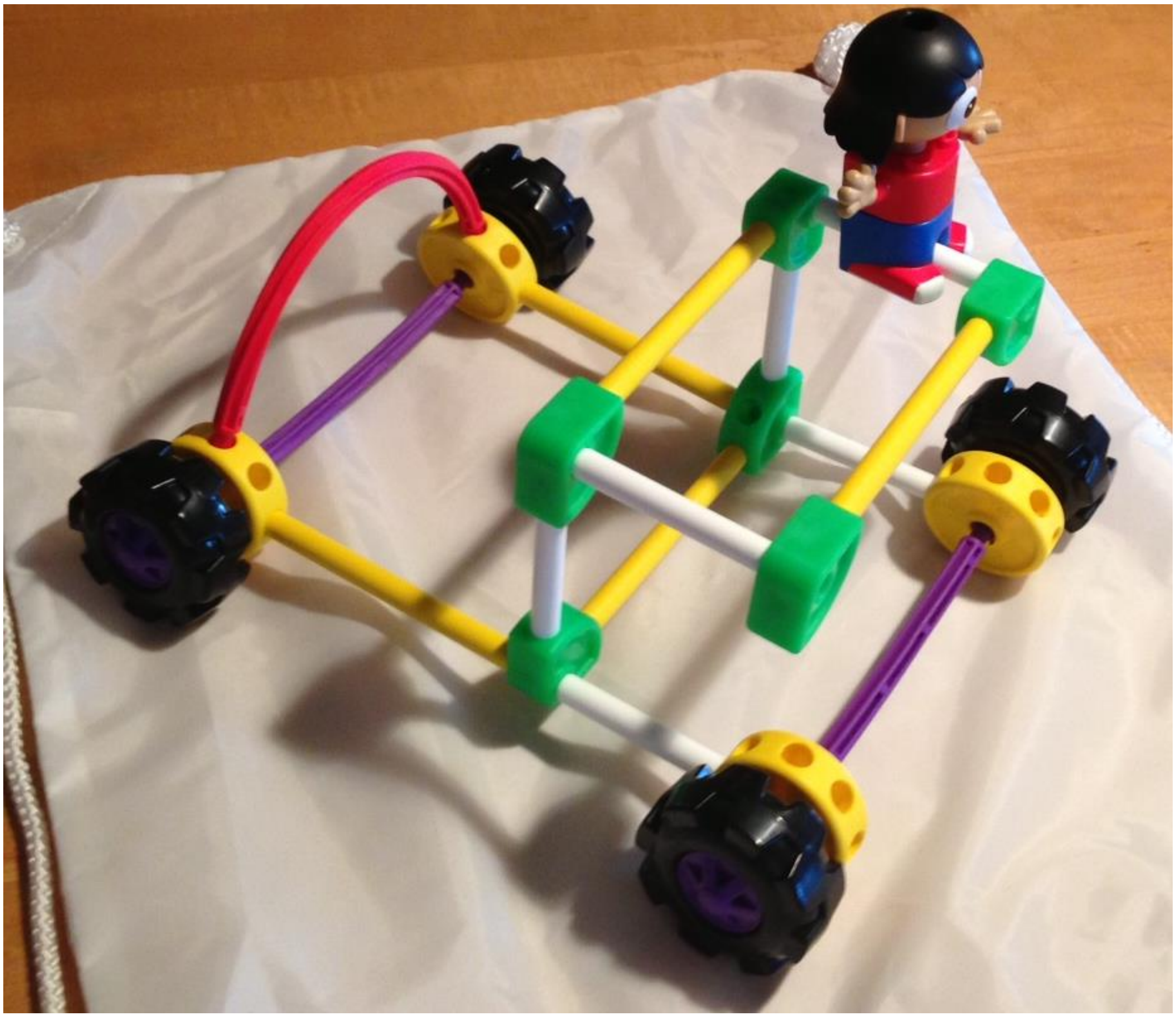
Score: time to complete



# APPENDIX E

Difficult Transfer Task

| Score | Code |
| --- | --- |
| 1 | Unfinished and includes 1+ errors |
| 2 | Unfinished but correct so far |
| 3 | Finished but includes 1+ errors |
| 4 | Finished correctly |



**APPENDIX F**

STEM Toys Rated for Future Play

| | |
|---|---|
| **Feminine Toys** | <br>*Rainbow Loom* |
| <br>*Roominate* | <br>*Pink Magnext* |
| <br>*Pink Tinker Toys (with girl)* | <br>*LEGO Friends* |

**Neutral**


*Blocks*

*Tangrams*


*Lincoln Logs*


*Puzzle*


*Toober & Zots*

<table>
<tr>
<td>Masculine</td>
<td>

Pirate K'Nex</td>
</tr>
<tr>
<td>
Masculine Magnext</td>
<td>

Super Hero LEGOs</td>
</tr>
<tr>
<td>

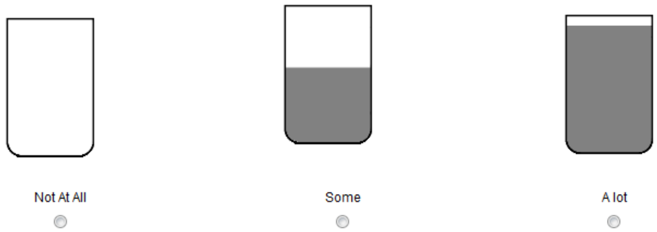
Vehicle Tinker Toys (with boy)</td>
<td>

Chemistry Set</td>
</tr>
</table>

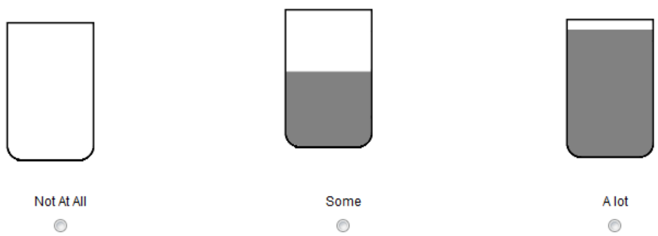
# APPENDIX G

Perceived Similarity to Goldie/Bobby Character Scale

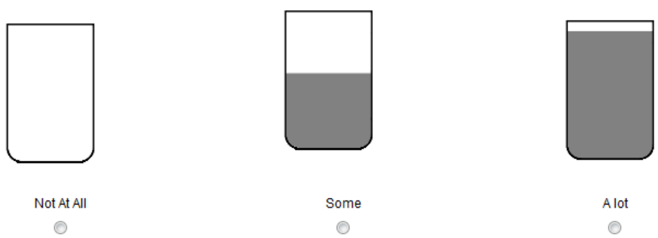
1. Think about how [character] looks. How much are you like that?



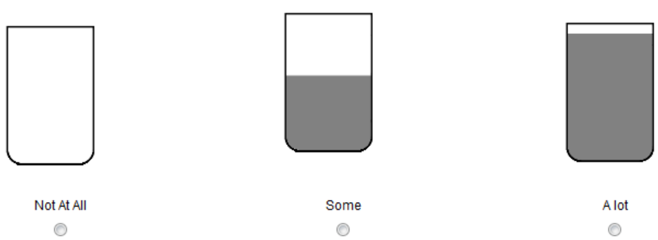

2. Think about the things [character] does. How much are you like that?



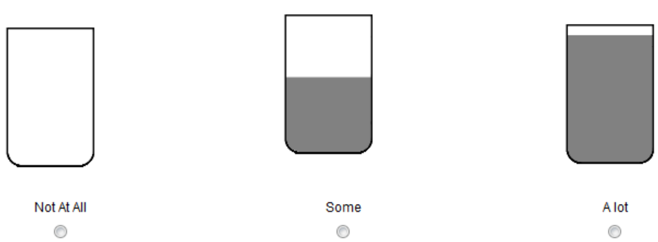

3. Think about the way [character] dresses. How much are you like that?




4. Think about the way [character] acts. How much are you like that?

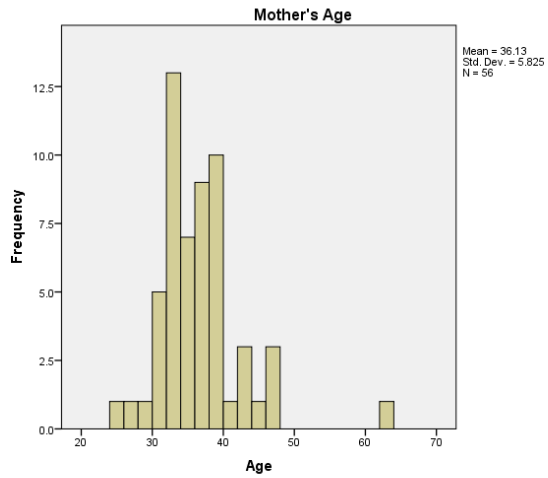


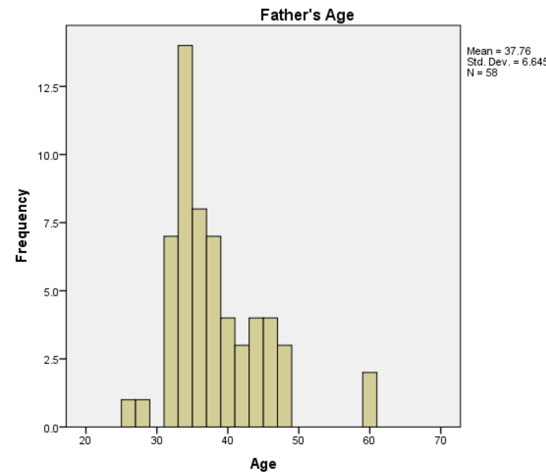

5. Think about what [character] is good at. How much are you like that?

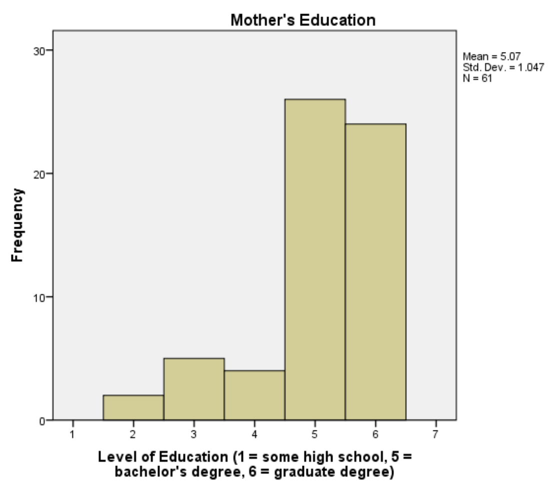


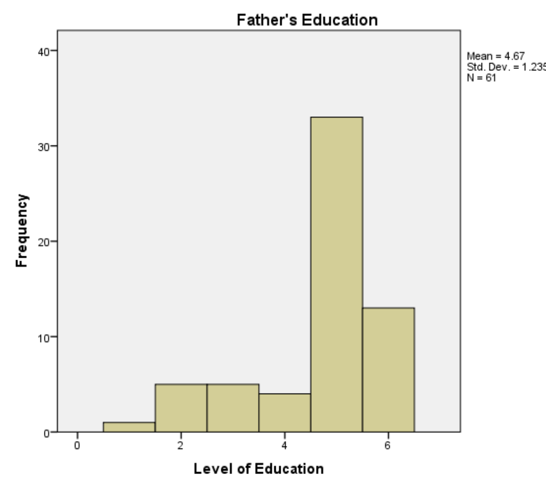
# Appendix H

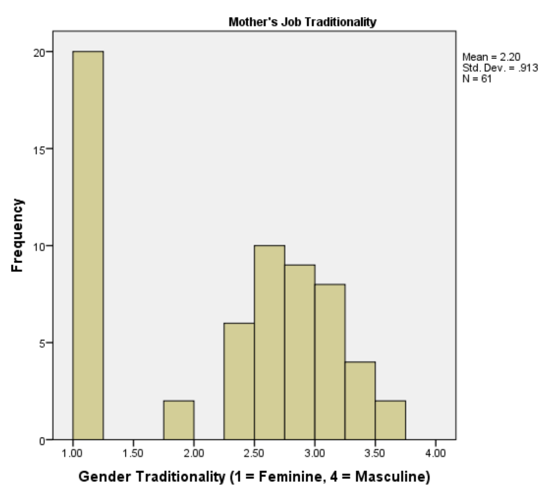
## Histograms of Family Characteristics

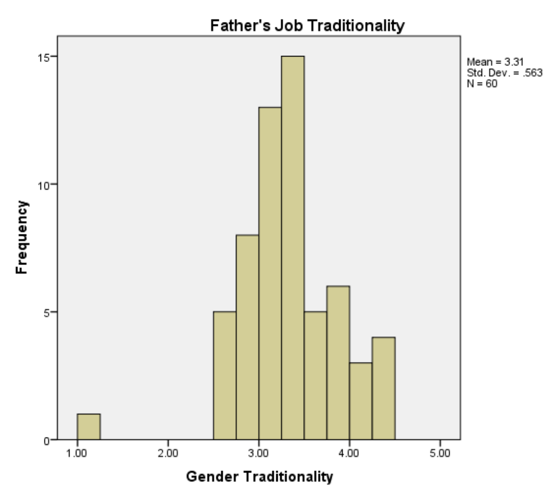


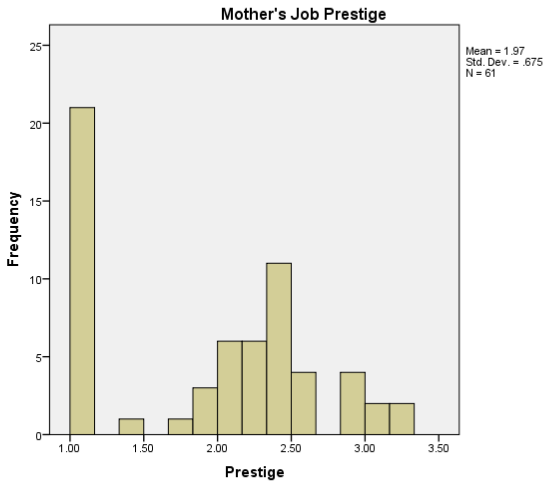



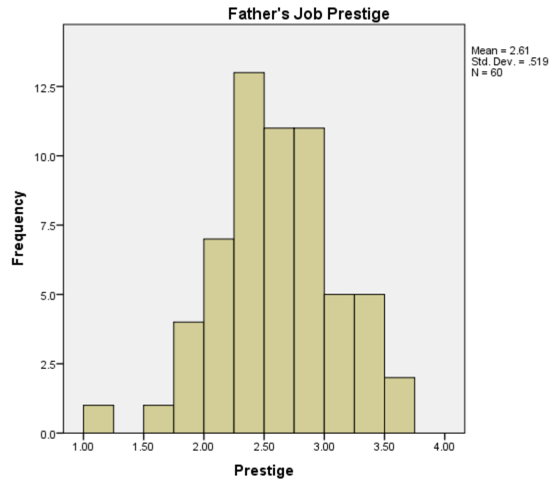



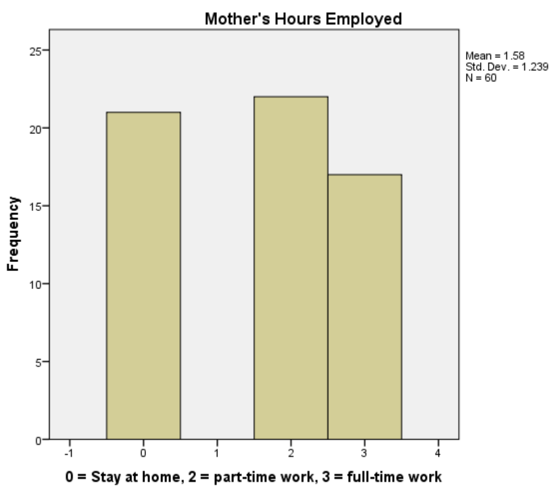



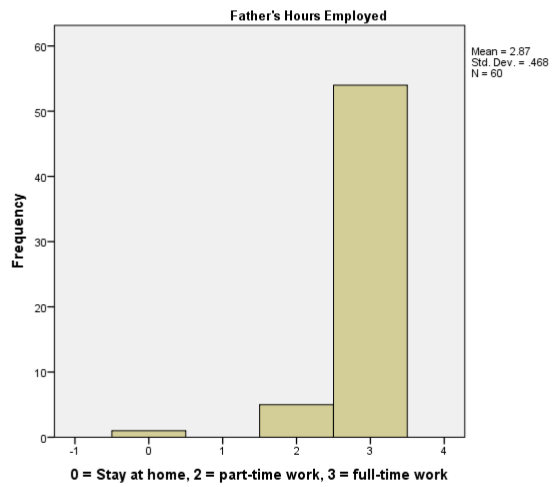



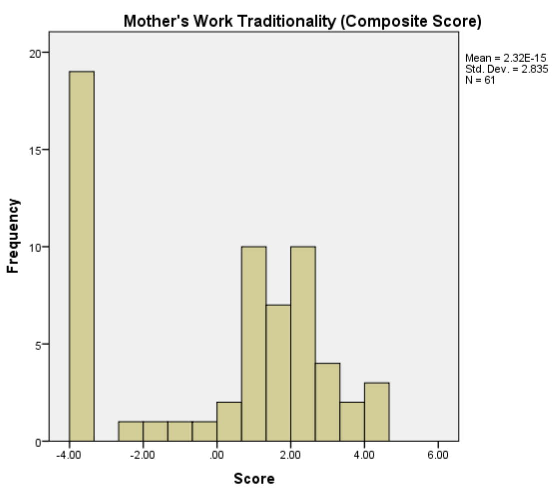



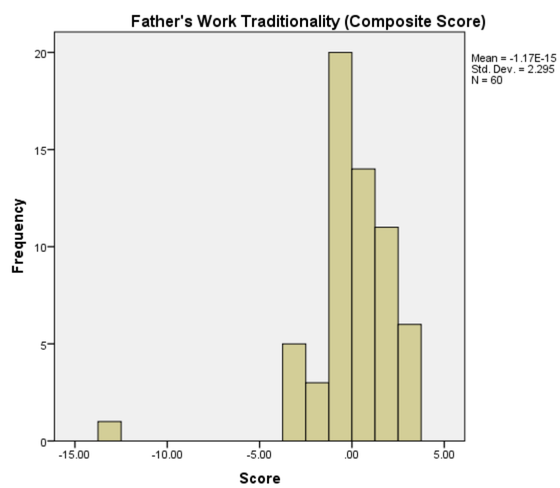

Mother's Job Prestige
Mean = 1.97
Std. Dev. = .675
N = 61
Frequency
Prestige

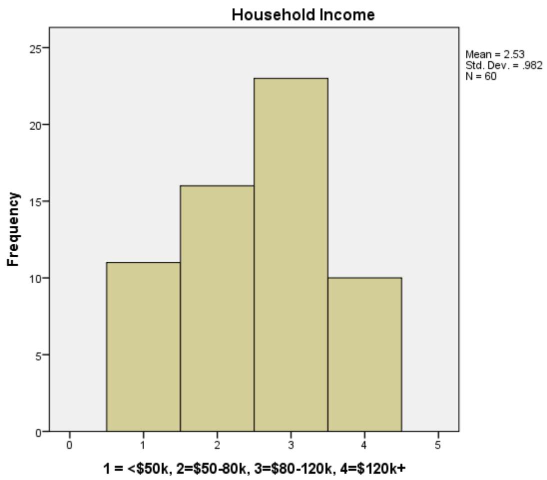

Father's Job Prestige
Mean = 2.61
Std. Dev. = .519
N = 60
Frequency
Prestige

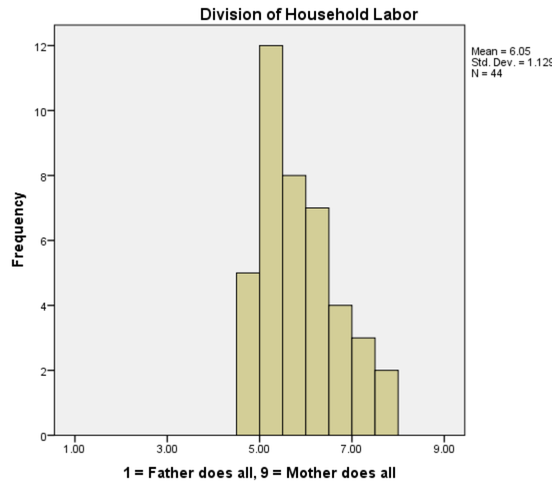

Mother's Hours Employed
Mean = 1.58
Std. Dev. = 1.239
N = 60
Frequency
0 = Stay at home, 2 = part-time work, 3 = full-time work

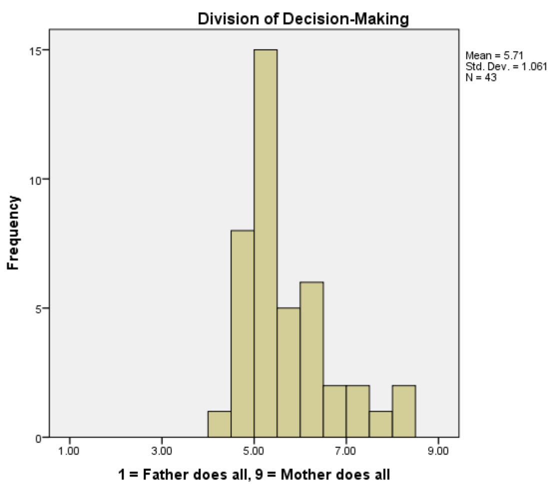

Father's Hours Employed
Mean = 2.87
Std. Dev. = .468
N = 60
Frequency
0 = Stay at home, 2 = part-time work, 3 = full-time work

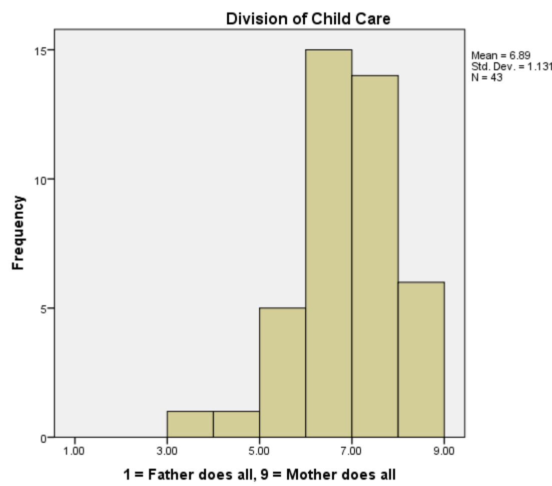

Mother's Work Traditionality (Composite Score)
Mean = 2.32E-15
Std. Dev. = 2.835
N = 61
Frequency
Score


Father's Work Traditionality (Composite Score)
Mean = -1.17E-15
Std. Dev. = 2.295
N = 60
Frequency
Score


Household Income
Mean = 2.53
Std. Dev. = .982
N = 60
Frequency
1 = <$50k, 2=$50-80k, 3=$80-120k, 4=$120k+

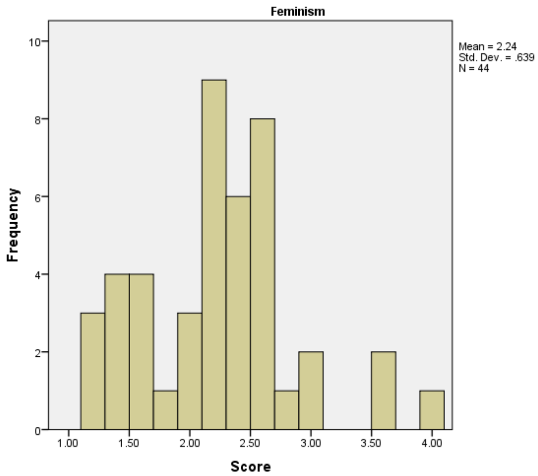

Division of Household Labor
Mean = 6.05
Std. Dev. = 1.129
N = 44
Frequency
1 = Father does all, 9 = Mother does all

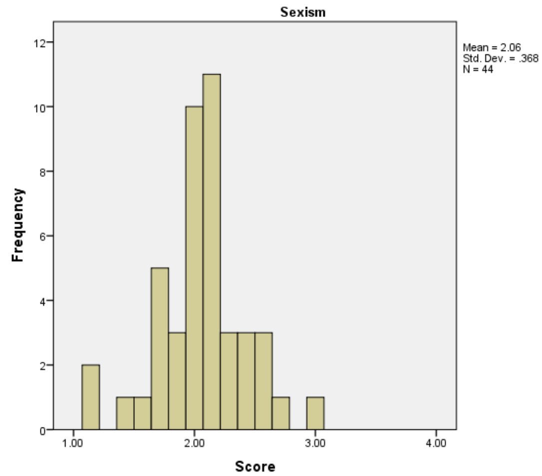

Division of Decision-Making
Mean = 5.71
Std. Dev. = 1.061
N = 43
Frequency
1 = Father does all, 9 = Mother does all

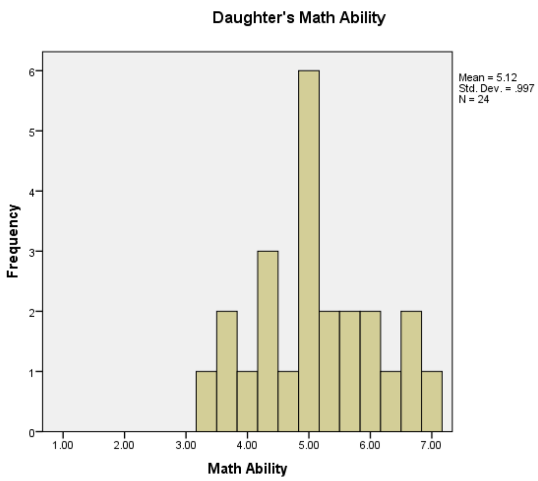

Division of Child Care
Mean = 6.89
Std. Dev. = 1.131
N = 43
Frequency
1 = Father does all, 9 = Mother does all

# Appendix I

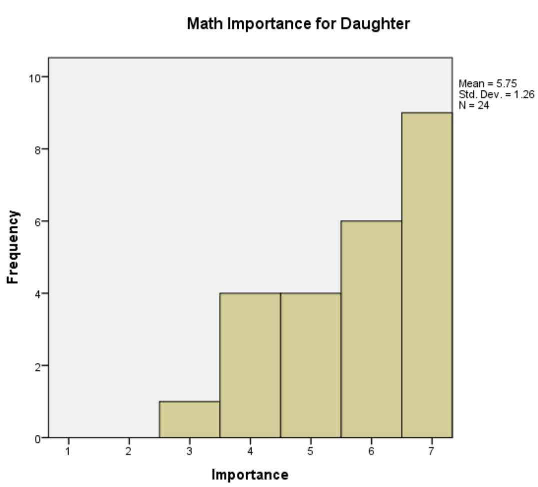
Histograms of Selected Individual Difference Measures for Mothers

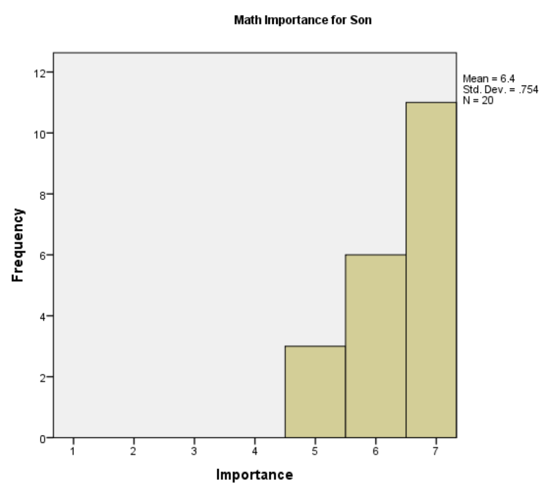


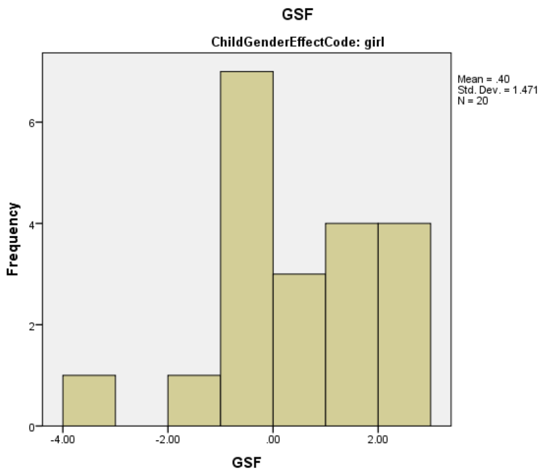


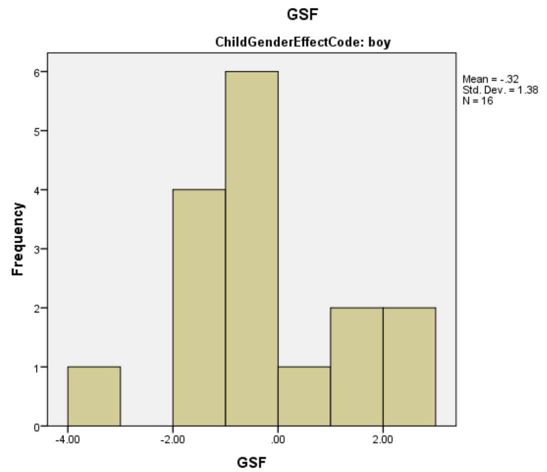


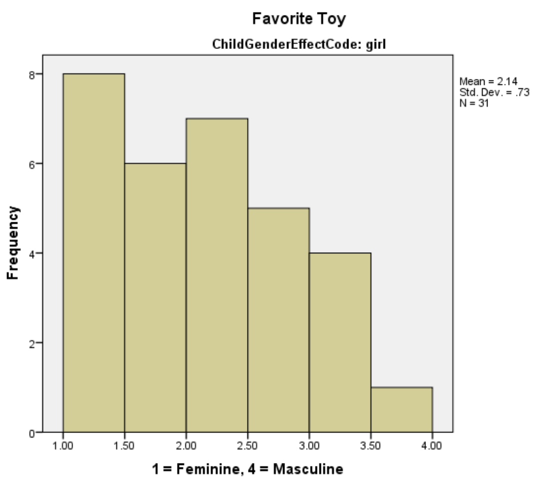



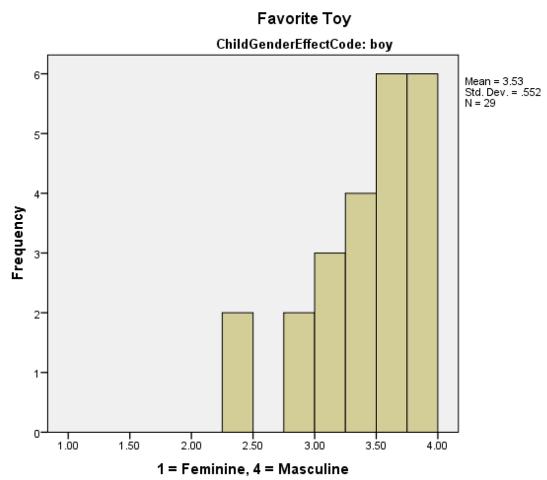
# Appendix J

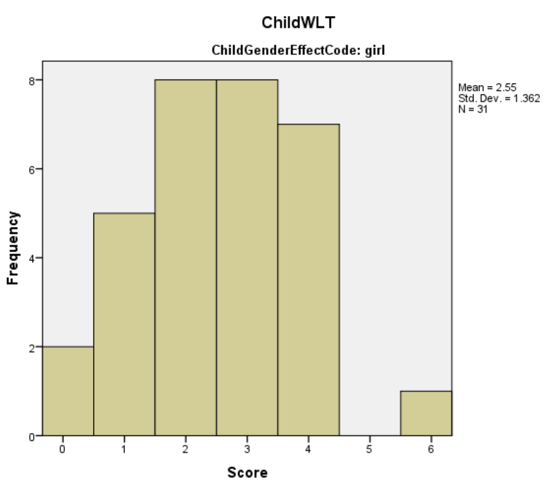
Histograms of Selected Individual Difference Measures for Children, Split By Gender

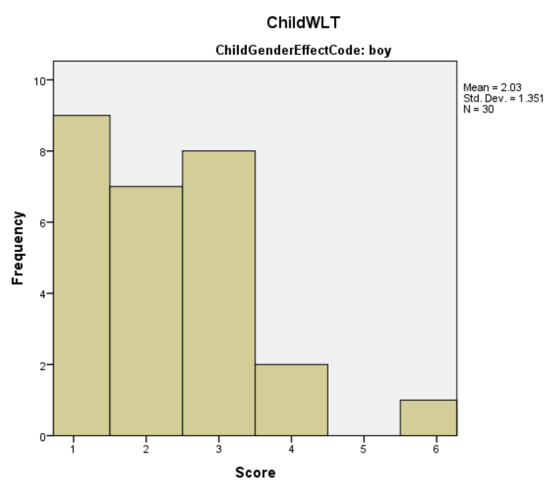


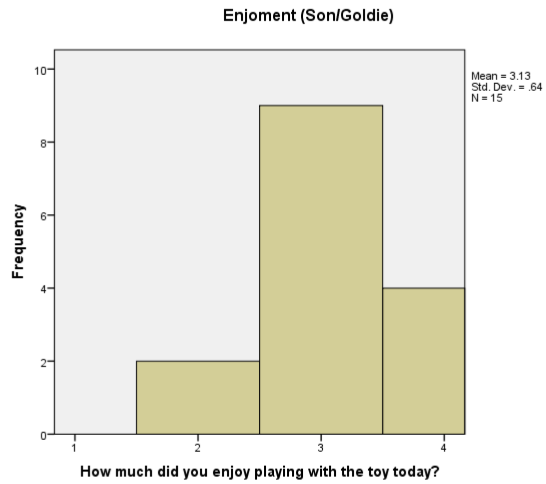



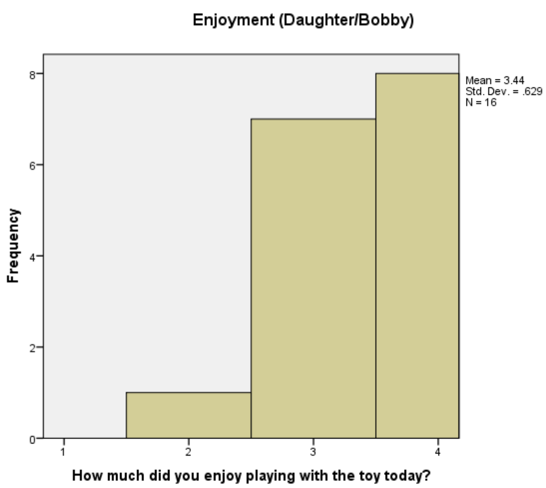



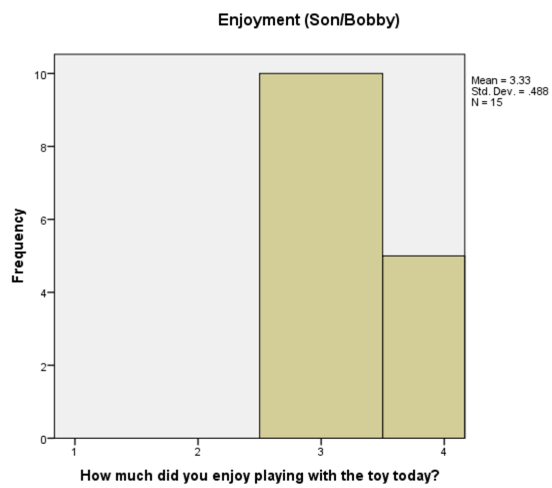



# Appendix K

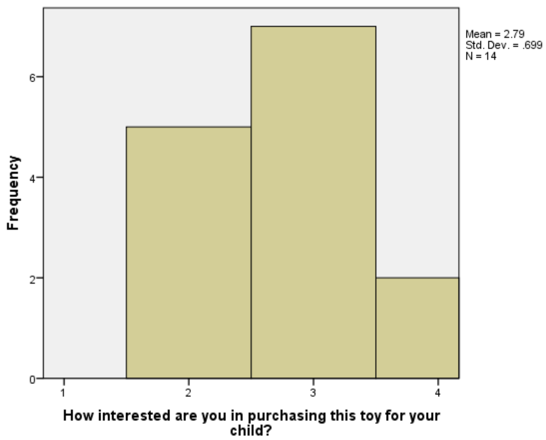
Histograms of Mothers' Outcome Measures, Split by Child Gender and/or Condition

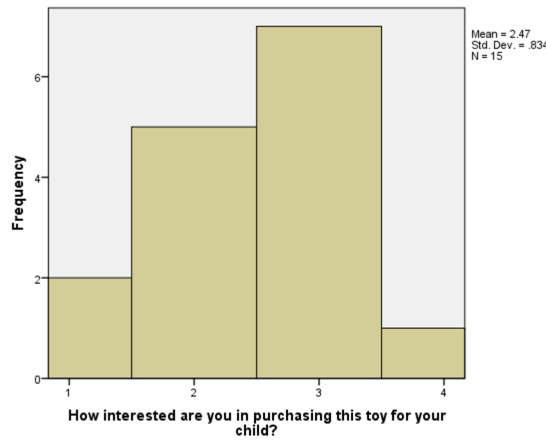



Interest in Purchase (Daughter/Goldie)

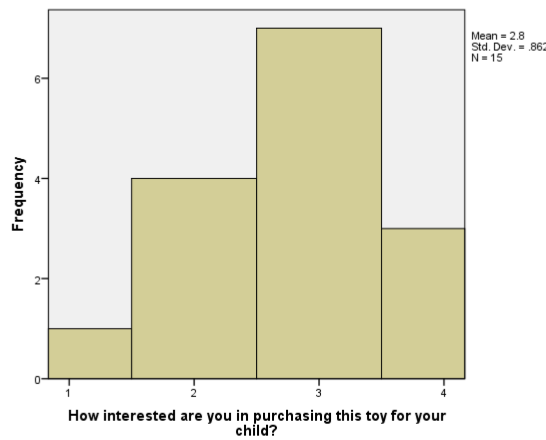


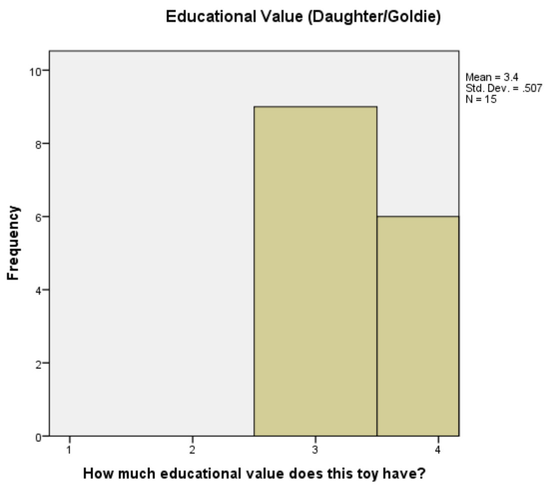

Interest in Purchase (Son/Goldie)

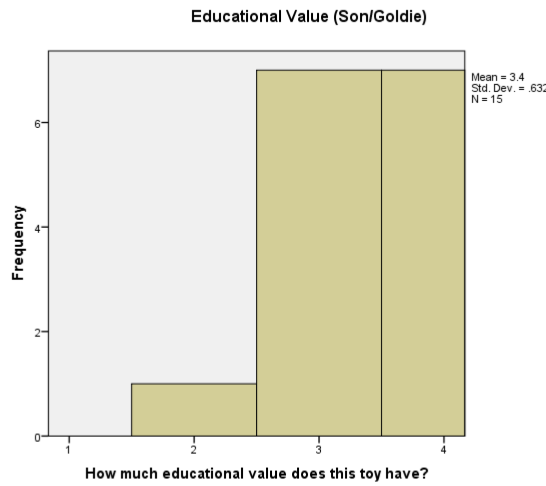


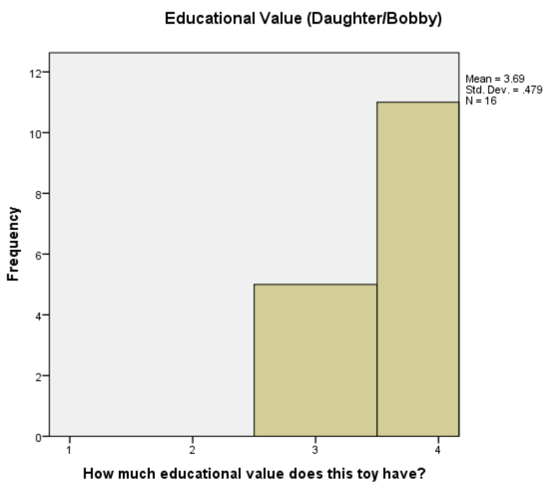

Interest in Purchase (Daughter/Bobby)

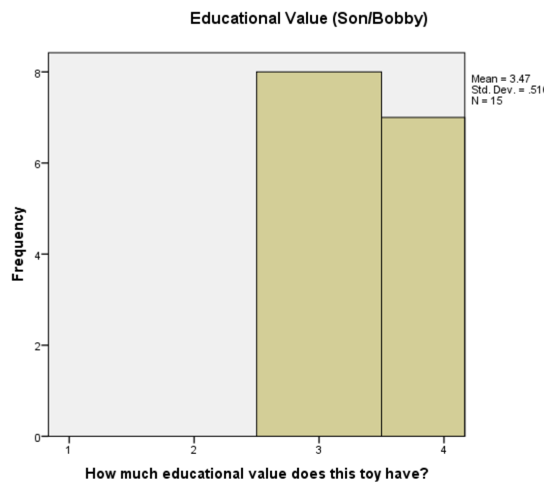

Interest in Purchase (Son/Bobby)

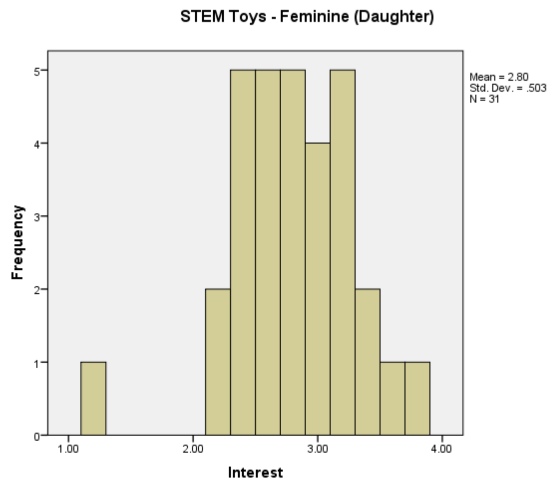


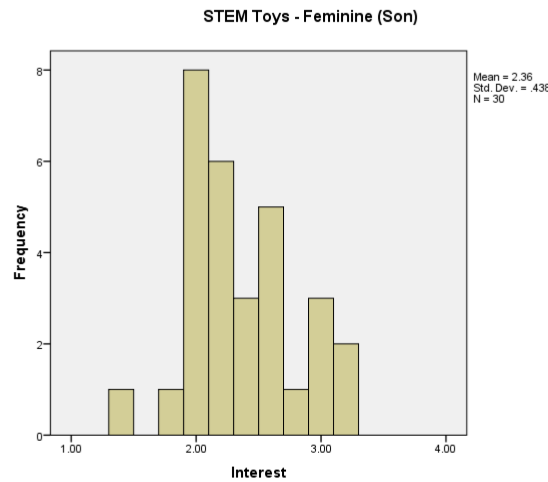

Educational Value (Daughter/Goldie)
Mean = 3.4
Std. Dev. = .507
N = 15
Frequency
How much educational value does this toy have?

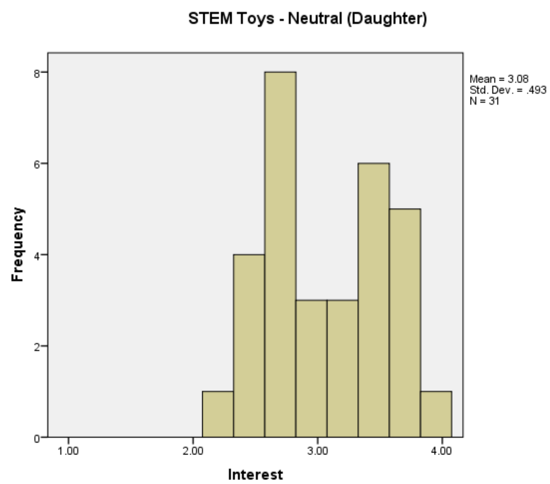

Educational Value (Son/Goldie)
Mean = 3.4
Std. Dev. = .632
N = 15
Frequency
How much educational value does this toy have?

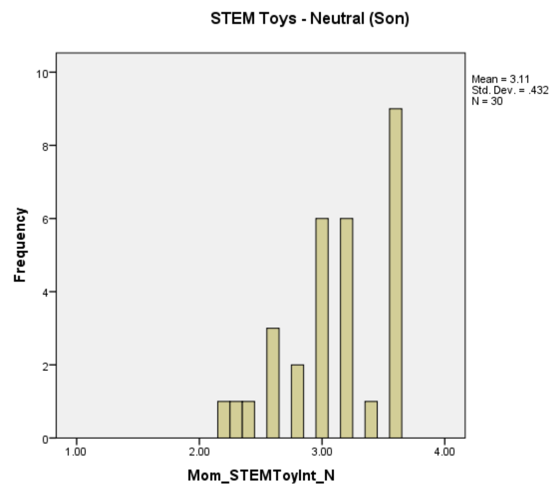

Educational Value (Daughter/Bobby)
Mean = 3.69
Std. Dev. = .479
N = 16
Frequency
How much educational value does this toy have?

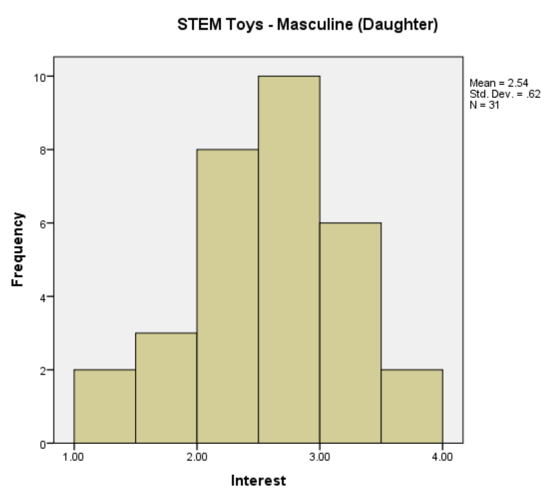

Educational Value (Son/Bobby)
Mean = 3.47
Std. Dev. = .516
N = 15
Frequency
How much educational value does this toy have?

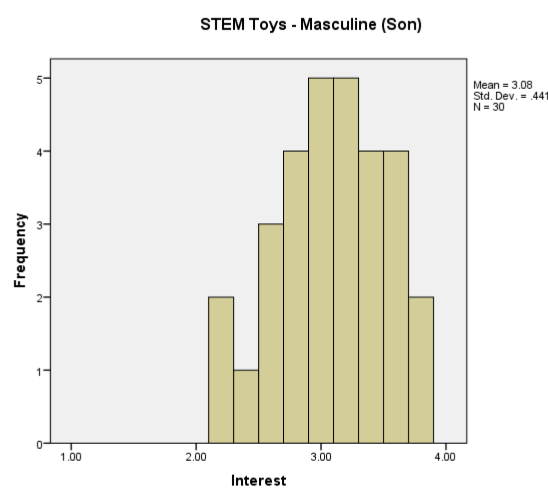

STEM Toys - Feminine (Daughter)
Mean = 2.80
Std. Dev. = .503
N = 31
Frequency
Interest


STEM Toys - Feminine (Son)
Mean = 2.36
Std. Dev. = .438
N = 30
Frequency
Interest


STEM Toys - Neutral (Daughter)
Mean = 3.08
Std. Dev. = .493
N = 31
Frequency
Interest

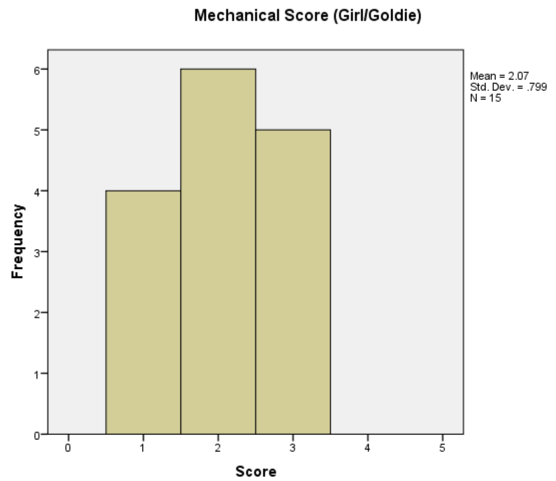

STEM Toys - Neutral (Son)
Mean = 3.11
Std. Dev. = .432
N = 30
Frequency
Mom_STEMToyInt_N

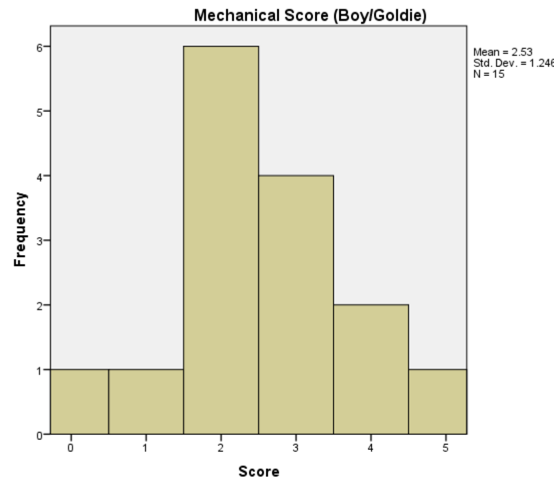

STEM Toys - Masculine (Daughter)
Mean = 2.54
Std. Dev. = .62
N = 31
Frequency
Interest

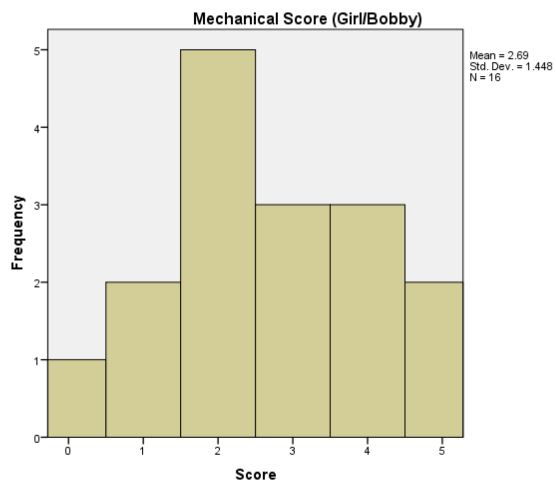

STEM Toys - Masculine (Son)
Mean = 3.08
Std. Dev. = .441
N = 30
Frequency
Interest

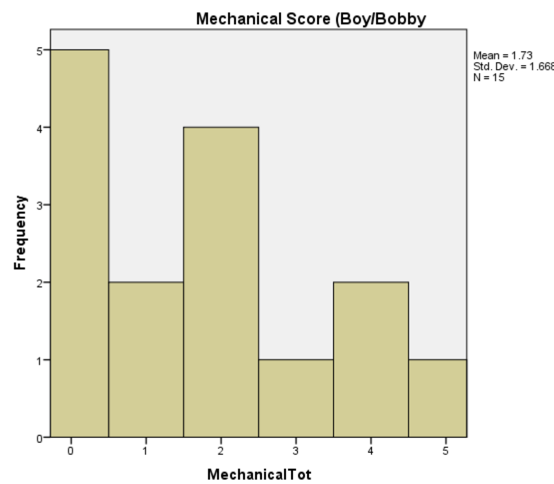
# Appendix L

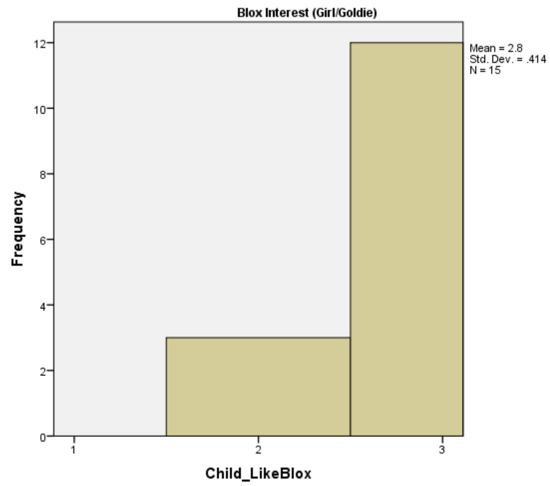
Histograms of Selected Child Outcome Measures, Split By Gender and/or Condition

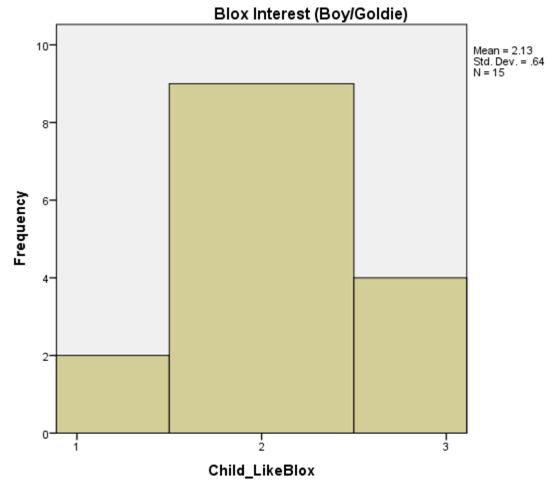


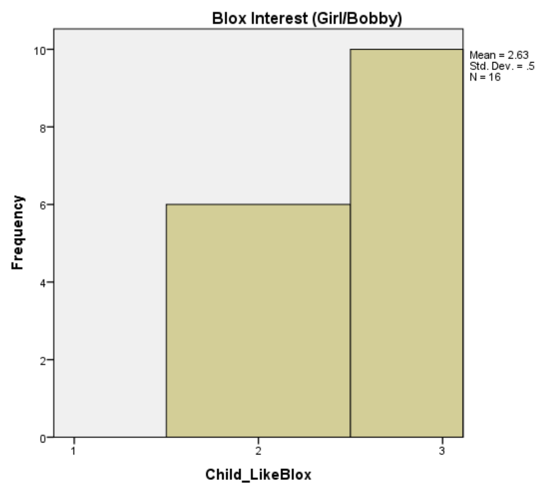



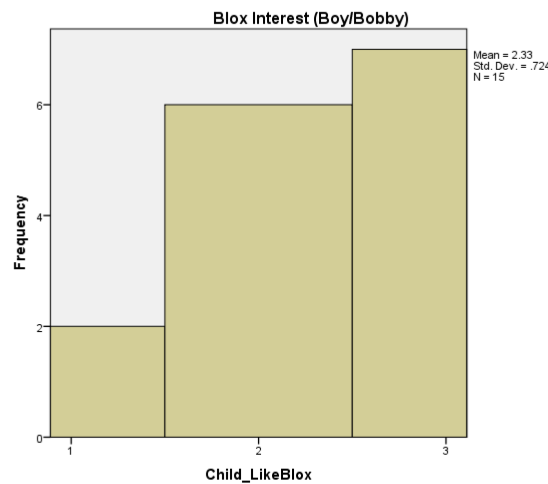



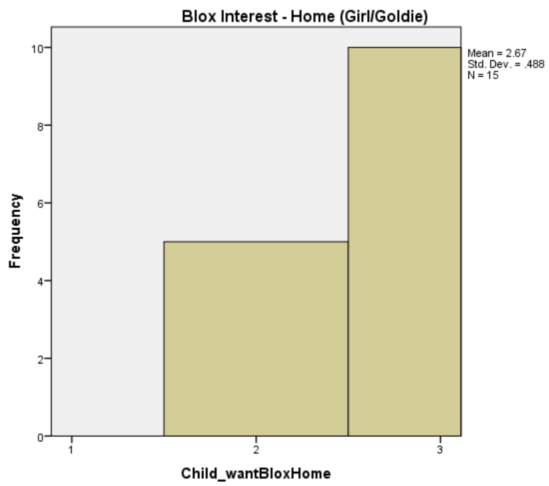



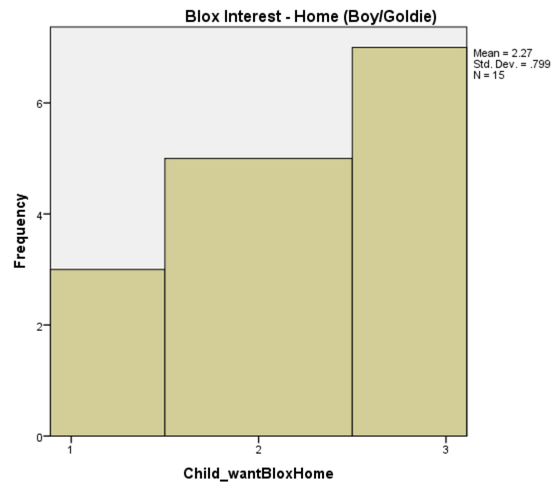

Blox Interest (Girl/Goldie)
Mean = 2.8
Std. Dev. = .414
N = 15
Frequency
Child_LikeBlox

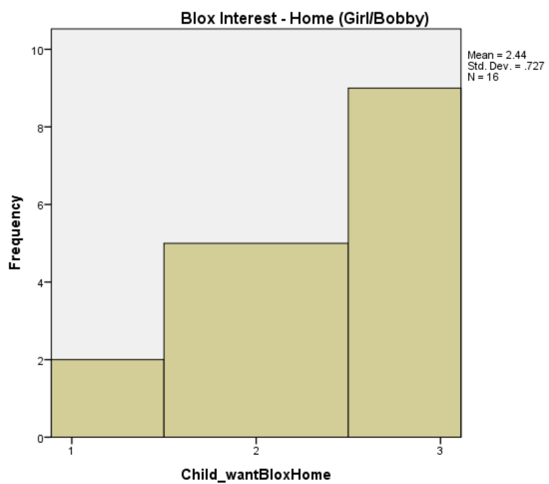

Blox Interest (Boy/Goldie)
Mean = 2.13
Std. Dev. = .64
N = 15
Frequency
Child_LikeBlox

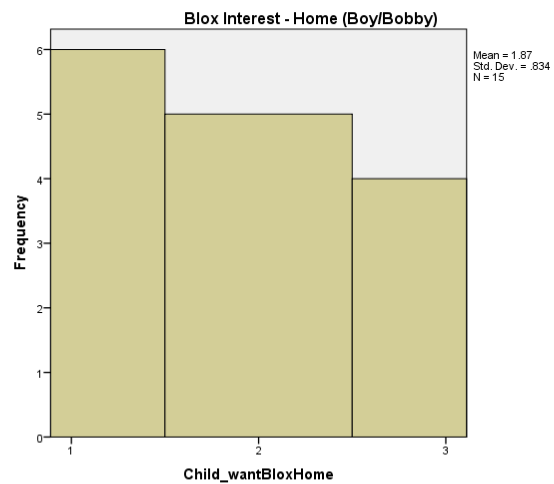

Blox Interest (Girl/Bobby)
Mean = 2.63
Std. Dev. = .5
N = 16
Frequency
Child_LikeBlox

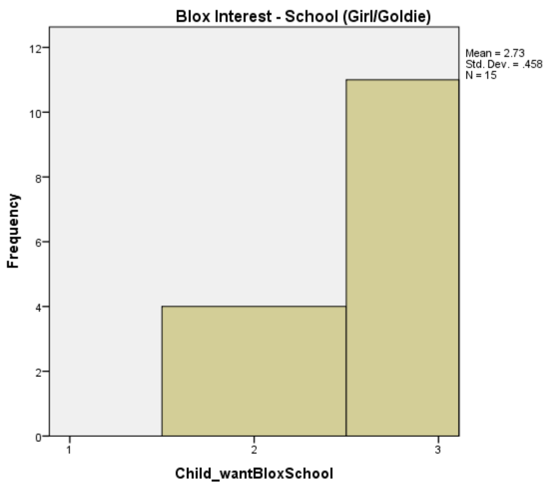

Blox Interest (Boy/Bobby)
Mean = 2.33
Std. Dev. = .724
N = 15
Frequency
Child_LikeBlox

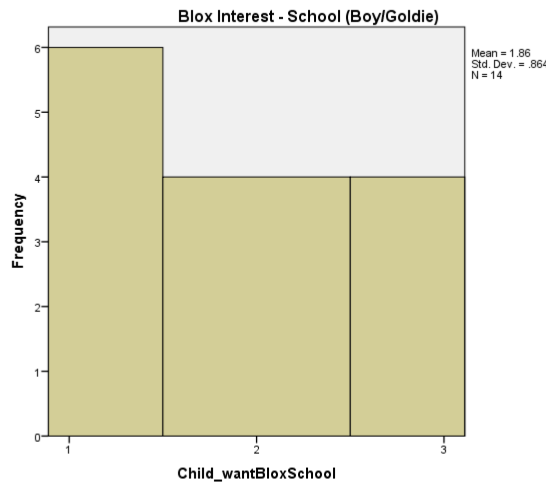

Blox Interest - Home (Girl/Goldie)
Mean = 2.67
Std. Dev. = .488
N = 15
Frequency
Child_wantBloxHome

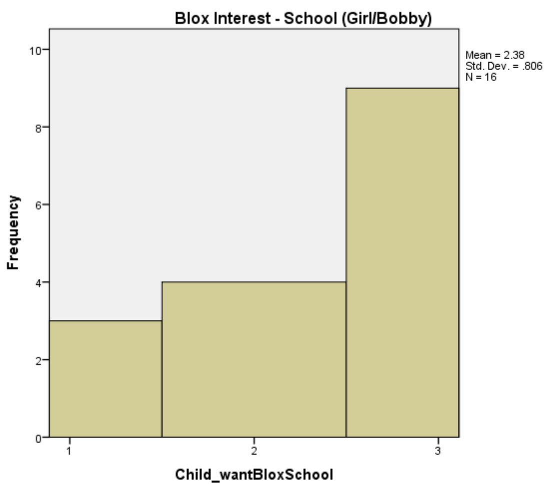

Blox Interest - Home (Boy/Goldie)
Mean = 2.27
Std. Dev. = .799
N = 15
Frequency
Child_wantBloxHome


Blox Interest - Home (Girl/Bobby)
Mean = 2.44
Std. Dev. = .727
N = 16
Frequency
Child_wantBloxHome

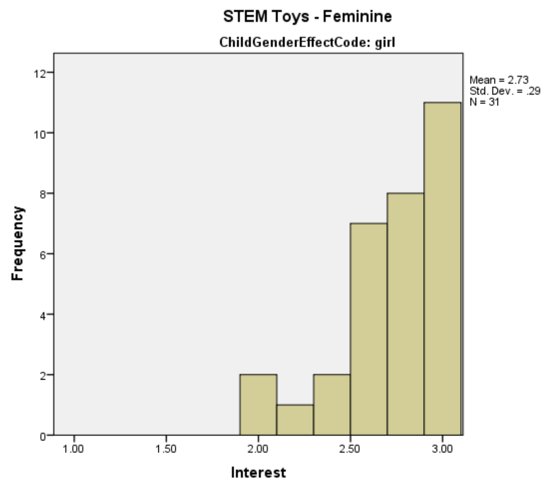

Blox Interest - Home (Boy/Bobby)
Mean = 1.87
Std. Dev. = .834
N = 15
Frequency
Child_wantBloxHome

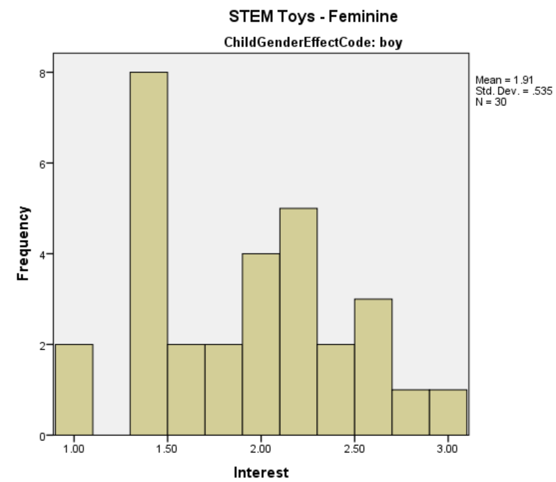

Blox Interest - School (Girl/Goldie)
Mean = 2.73
Std. Dev. = .458
N = 15
Frequency
Child_wantBloxSchool

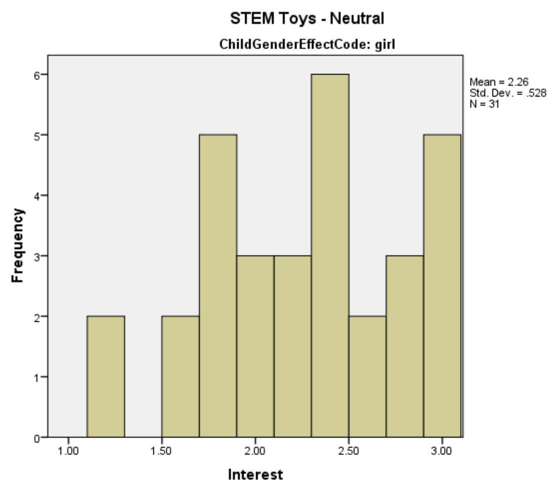

Blox Interest - School (Boy/Goldie)
Mean = 1.86
Std. Dev. = .864
N = 14
Frequency
Child_wantBloxSchool

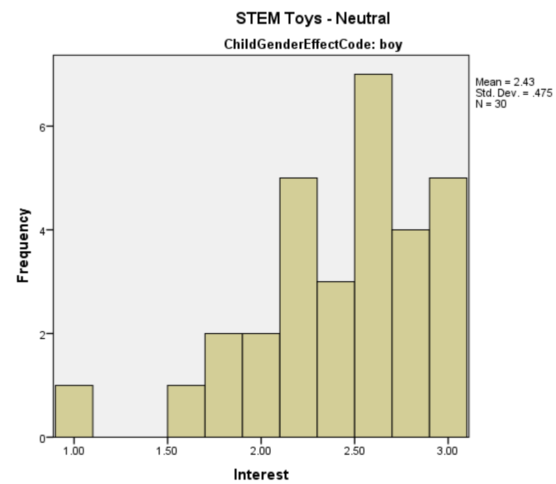

Blox Interest - School (Girl/Bobby)
Mean = 2.38
Std. Dev. = .806
N = 16
Frequency
Child_wantBloxSchool

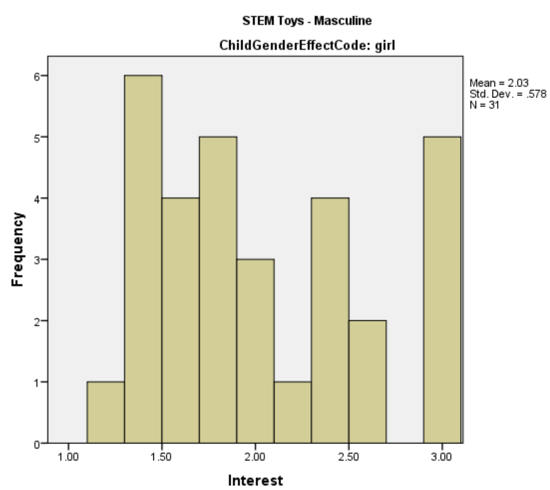

Blox Interest - School (Boy/Bobby)
Mean = 2.13
Std. Dev. = .834
N = 15
Frequency
Child_wantBloxSchool

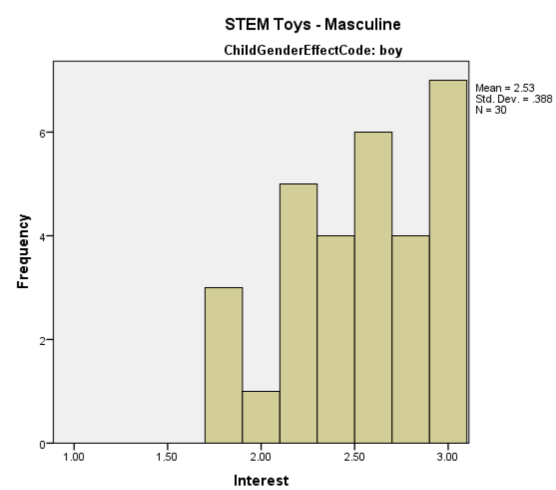

STEM Toys - Feminine
ChildGenderEffectCode: girl
Mean = 2.73
Std. Dev. = .29
N = 31
Frequency
Interest


STEM Toys - Feminine
ChildGenderEffectCode: boy
Mean = 1.91
Std. Dev. = .535
N = 30
Frequency
Interest


STEM Toys - Neutral
ChildGenderEffectCode: girl
Mean = 2.26
Std. Dev. = .528
N = 31
Frequency
Interest


STEM Toys - Neutral
ChildGenderEffectCode: boy
Mean = 2.43
Std. Dev. = .475
N = 30
Frequency
Interest


STEM Toys - Masculine
ChildGenderEffectCode: girl
Mean = 2.03
Std. Dev. = .578
N = 31
Frequency
Interest


STEM Toys - Masculine
ChildGenderEffectCode: boy
Mean = 2.53
Std. Dev. = .388
N = 30
Frequency
Interest

**Emily F. Coyle**

**Curriculum Vitae**

emilyfcoyle@gmail.com
http://emilyfcoyle.wix.com/emilyfcoyle

## Education

**Ph.D.,** Psychology, The Pennsylvania State University, 2015

**M. S.,** Psychology, The Pennsylvania State University, 2012

**B. S.,** Psychology, Summa Cum Laude, Washington and Lee University, 2010

## Awards and Grants

**National Science Foundation Graduate Research Fellowship**, 2011

**University Graduate Fellowship**, Penn State, 2010

**Graduate Exhibition Award**, First Place in Social & Behavioral Sciences, 2014

**RGSO Dissertation Award**, College of the Liberal Arts, 2013

**Dissertation Award**, Child Study Center's Robert & Ruth Faris Child Psychology Fund, 2013

**Morgan Scholarship**, College of the Liberal Arts, 2010

**Student Travel Award**, Society for Research in Child Development, 2015

**Travel Awards**, Penn State Child Study Center 2011-2015

**Phi Beta Kappa**, inducted Winter 2010

## Publications

**Coyle, E. F.**, Van Leer, E., Fulcher, M., & Schroeder, K. M. (in press). Planning to have it all: Emerging adults' expectations for future work-family conflict. *Sex Roles.*

Liben, L. S. & **Coyle, E. F.** (2014). Developmental interventions to address the STEM gender gap: Exploring intended and unintended consequences. In L. S. Liben & R. S. Bigler (Eds.), *The role of gender in educational contexts and outcomes*. In J. Benson (Series Ed.), *Advances in child development and behavior* (Vol. 47, pp. 77-116). London: Elsevier.

Fulcher, M., & **Coyle, E. F.** (2011). Breadwinner and caregiver: A cross-sectional analysis of children's and emerging adults' visions of their future family roles. *British Journal of Developmental Psychology, 29*, 330-346. doi:10.1111/j.2044-835X.2011.02026.x

## Research Positions

**Graduate Research Fellow**, Cognitive and Social Development Lab, Dr. Lynn Liben, The Pennsylvania State University, August 2010 – Spring 2015

**Research Assistant**, Cognitive Aging and Neuroimaging Lab, Dr. Nancy Dennis, The Pennsylvania State University, Summer 2011 – Spring 2013

## Teaching Experience

**Instructor**, The Pennsylvania State University, PSYCH 200: Elementary Statistics in Psychology, Summer 2014, Fall 2014, Spring 2015 (Instructor, online World Campus)

**Instructor**, The Pennsylvania State University, PSYCH 479: Psychology of Gender, Spring 2013, Fall 2013 (Independent Instructor)

**Schreyer Institute for Teaching Excellence Graduate Teaching Certificate**, 2014

**Penn State World Campus Certificate for Online Teaching**, 2014