Children’s Growing Understanding of Personal Identity: Gender Stereotypes and Play

Riley Abeles

Advisor: Beth Hennessey

Submitted in Partial Fulfillment of the Prerequisite for Honors in Psychology
Wellesley College

April 2016
Acknowledgements

My thesis would not be what it is without the support of so many individuals.

I would like to thank Becky Geer and Annie Cohen for their continued feedback and support at every stage of my thesis. Jennifer Clifton, Eli Cohen, Wendy Gosenhauser, and Emily Diesl, thank you for your help and support throughout the experimental process. Susan Ferguson and Hanna Chen, thank you for providing a wonderful learning environment. I am so grateful for having had the chance to learn from you all.

Thank you to Linda Carli, Sally Theran, and Sarah Wall-Randell for serving on my thesis committee and for your feedback and suggestions this year. Thank you also to Tracy Gleason, Margaret Keane, Jenny Pyers, and Christen Deveney for your encouragement throughout the process. Additionally, thank you to Hea Jung Lee and Camille Stovall-Ceja for all of your work in data collection and the coding process. Thank you to the Dean of the College for the financial support of my thesis.

I am so glad for having had the support of Raizel DeWitt, Grace Bennett-Pierre, Jenny Chen, Ali Roule, and Marissa Klee-Peregon, as we have worked through our theses together. Thank you to my family and friends for your support and feedback this year.

I would especially like to acknowledge Beth Hennessey for being such an enthusiastic and committed advisor. I am so appreciative of all of the time you have put in to reading and editing drafts, helping me focus my ideas and plans, and working to design a study I am proud of. Your support and unwavering interest in my research has helped shape my thesis from a vague idea into an actual research paper and I am very grateful for having had the chance to work with you.
Abstract

Gender is one of the more complicated identity traits to model, investigate, and comprehend. Despite this complexity, researchers have long found that the establishment of gender identity is fundamental to even a very young child’s developing sense of self. As children strive to learn about complex social norms, including those regarding gender, play is one of their best tools for comprehending and exploring social rules and expectations. Through play, children are able to express and process their emotions, and in interaction with their peers, they are also given the opportunity to recreate and explore social dynamics (Kyratzis & Ervin-Tripp, 1999; Russ & Fiorelli, 2010). In fantasy play, the themes children explore are strongly related to their own real-world behaviors and home lives (Von Klitzing, Kelsay, Emde, Robinson, & Schmitz, 2000). In the present study, the relations between children’s formation of and adherence to gender stereotypes, their preferences for specific styles of play, their playmates, and the narratives they create during play were explored in a laboratory preschool setting. Results showed both that these preschoolers were well aware of prevailing gender stereotypes, and that these stereotypes were frequently exhibited in their own gendered behavior, as displayed during naturalistic play observations and an experimenter-designed play situation. Data from this 2016 investigation are compared to results reported in previous studies carried out two or more decades ago and similarities and differences across experimental findings are discussed.
Children’s Growing Understanding of Personal Identity: Gender Stereotypes and Play

Gender Schemas

Children are not born with an innate understanding of the world around them. They must develop that understanding over time through a variety of complex processes. Children’s increasing levels of cognitive sophistication as they grow and develop allow for comprehension of progressively complicated constructs and relations in the world. Gender is one of the more complex identity traits to model, investigate, and comprehend. Physiology, clothing, behavior, interests, and personality each contribute to a child’s developing gender identity, as do cultural values and societal norms. Despite this complexity, researchers have long found that the establishment of gender identity is fundamental to even a very young child’s developing sense of self. When children are asked to describe themselves using only one trait, they tend to choose gender over other salient descriptors such as race, ethnicity, and age (Fischer, Hand, Watson, Van Parys & Tucker, 1984). By the age of 2, most typically developing children understand gender differences and identify themselves and others by gender (Ruble et al., 2007). By age 3, most children show evidence of gender stereotyping, and they demonstrate rigid gender stereotypes in their thinking by age 5 (Fischer et al., 1984).

Children employ developmentally appropriate cognitive strategies for information processing, and at a young age this means that much of their understanding of the world around them is formulated in terms of simple, concrete structures. Preschool-aged children’s conceptions about gender tend to be extremely concrete and are driven, at least in part, by the creation and elaboration of schemas. Schemas are information processing strategies that allow for the interpretation and processing of new information (Bem, 1983; Martin & Halverson, 1981). Schemas provide meaning to new input in that they allow for information to be grouped
GENDER STEREOTYPES AND PLAY

This organizational structure that schemas provide enables information to be processed and accessed more efficiently, and also allows for missing information in a new situation to be replaced with schematic knowledge over time (Martin & Halverson, 1981). Young children’s schemas are rigid and inflexible at their creation, and become more nuanced with age and time. For example, a child might initially create a schema for dogs based on her family’s pet chihuahua. This schema might involve the information that a dog has four legs and is small with a shrill bark. When the child’s mother points out a golden retriever while on a walk, saying, “Look, a dog!” the child’s schema will expand to involve this new information about dogs. She will learn that all dogs have four legs and are furry, but they can be large or small, and have a bark that is shrill or deep. As explained by Bem (1983), a definitional category such as “dog” will eventually become a schema if the social context attributes broad social significance to the category, and the child continues to have new experiences that lead to the association of the category with additional attributes.

Children’s understanding of gender evolves through a similar process of schema formation and modification. Gender is made salient to children as a grouping category in terms of its binary categorization, obvious physical markers, and universality (Barberá, 2003). Children’s early gender schemas tend to focus on physical markers, such as clothing and appearance, and later include stereotypically masculine or feminine items and activities—they learn that pink is for girls and blue for boys, but also that certain toys are more appropriate for boys or girls. As children are presented every day with new and different information about gender, they learn to categorize physical and social traits as masculine or feminine. This information is encoded and organized into their evolving gender schema (Bem, 1983). However, young children’s tendency to think concretely challenges their ability to understand the nuances
of gender. For adults, gender differences take on social significance for their role in reproduction and social norms, but children learn to define gender in terms of outwardly observable and obvious features (Bigler & Liben, 2007). In fact, it is not until approximately age 10 or 11 that most children even begin to understand the subtler gender markers and stereotypes (Martin & Halverson, 1981).

The development of a sense of self is crucial for both adults and children. Martin and Halverson (1981) suggest that children’s tendency to employ categorical processing strategies such as schemas not only allows them to process information more efficiently, but also encourages the development and understanding of a personal identity. Children must come to make sense of the world around them, and in addition, their place in it. Gender is frequently emphasized for young children in the context of their everyday social experiences, such as when a teacher addresses her class as “boys and girls,” or an adult praises the gender-stereotyped behaviors that a child exhibits (Bigler & Liben, 2007). The cultural significance of gender is made evident to children right from the start and becomes especially important as they construct their growing worldviews. Importantly, unlike other individual differences surrounding personal style, preferences or opinions, gender is a category that is perceptually salient—markers of gender are frequently physical and/or tied to other aspects of outward appearance (Bigler & Liben, 2007). These physical, outward cues, coupled with the social salience of gender, encourage children’s developing identification as a boy or a girl (Bem, 1983). In fact, children tend to identify with their gender group earlier than they do with other salient traits such as race, ethnicity, or age (Fischer et al., 1984). By age two, most typically developing children understand gender differences and can easily identify themselves and others within their gender schemas (Bem, 1983; Ruble et al., 2007).
Over time, children’s gender schemas become increasingly complex and begin to strongly influence their actions. Escalera (2009) argues that when children feel that gender is relevant or salient, it is especially likely to affect their behavior and speech. Children are frequently subjected to gender role norms in school and at home. TV shows targeted at children, presents from family members, comments or directives from parents, admonitions and instructions from peers who are also learning about gender—all of these influences and more—flood children with information regarding social expectations based on gender. Any toy store aisle will demonstrate this phenomenon, as toys are clearly separated by color and theme: there is the pink and purple section, filled with flowers and princess, and then there is the section filled with trucks, vehicles, and superhero masks, dominated by red, blue, and dark colors.

Stereotypical messages about what boys and girls should enjoy doing or wearing, even messages about what it means to be a boy or a girl, are everywhere. Over time, children learn to monitor their own behavior, predict others’ behavior, and understand the social expectations they face based on gender (Martin & Halverson, 1981). Child development research reveals that gender schemas increase in rigidity until the age of five, at which time they become quite rigid, at least temporarily. In fact, by age five, children’s behavior may be more strongly affected by their gender schemas than at any other point in their lives (Ruble et al., 2007). For example, five-year-olds tend to sort toys according to gender stereotypes significantly more often than do three-year-olds (Freeman, 2007). And in a review of research focused on children’s gender stereotypes, Signorella, Bigler, and Liben (1993) found in the majority of cases that older children (ages 6 to 13) tended to give more gender stereotyped answers to questions about occupations, traits, and toys than did younger children (ages 2 to 3). The fact that emphasis on gender stereotypes increases with age might seem counterintuitive. At issue here, however, is the fact that children
are working hard to understand their own gender identity in the context of ever broadening and complex social situations. In order to accomplish this goal, they must come to understand that gender is fixed and internal across situations and contexts (Lewis & Brooks-Gunn, 2012). Children’s initial understanding of gender categories develops around the same time that they begin to understand object constancy—the idea that categories of objects have different traits that are persistent over time and across individual items (Lippa, 2002). The rigid thinking children display during this time period allows them to more easily solidify their understanding of gender constancy, the idea that a person’s gender identity is fixed and does not shift between contexts. For example, if a girl usually wears dresses, she needs to come to understand that she will not turn into a boy if she wears a pair of pants.

The rigidity of gender schemas is not consistent across children or even within children; children’s own traits and experiences greatly influence the schemas that they come to form, and schemas can therefore become more or less flexible over time. Gender schemas have differing levels of salience for boys and girls and affect their behavior differently. In analyzing the schematic knowledge that eleven-year-old children incorporated into their interpretations of a variety of hypothetical scenarios, Barberá (2003) found that girls were less likely than boys to rely on gender-stereotyped information. This difference is reflected in preschool-aged children as well, with research finding that young boys tend to use gender stereotypical information as a guide for sorting objects more frequently than do girls (Signorella et al., 1993). Additionally, boys tend to be more rigid in their formation of gender schemas for males than for females, and they also tend to demonstrate far more concern when males, rather than females, transgress gender stereotypes (Smetana, 1986). Elaborating on these and similar findings, Scott (1984)
suggests that girls may come to form more flexible gender schemas because in a patriarchal society, girls have more to gain from increased gender schema flexibility than do boys.

As their cognitive strategies develop further in later childhood, children become increasingly more likely to define themselves in terms of multiple identity traits; and, over time, the significance of gender as a defining trait is weakened (Martin & Halverson, 1981). Additionally, over time, children’s gender schemas become more flexible. This shift significantly influences children’s behavior and understanding of gender roles. Older children are better able than younger children to recall counter-stereotypical information from stories, such as a woman fixing a car or a man sewing clothes (Bigler & Liben, 1990). Importantly, however, the influence of young children’s early gender schemas must not be underestimated. Schema formation is an essential early cognitive strategy for young children to make sense of and learn about the environment and their place in it. Schemas make the processing of large amounts of information possible, but the rigidity of early gender schemas may prevent preschoolers from fully exploring certain opportunities. If a child’s behavior is limited by his gender schema, he might miss out on activities, games, or socialization and play experiences that could significantly enrich his developmental progression.

Play

Like the formation of schemas, play is another important vehicle for children to develop an understanding of the world and their place in it. These processes actually inform each other—as children learn about the world, they incorporate new knowledge into their play, which in turn allows them to more fully process and integrate new ideas into their worldview. Children’s play is difficult to define and the designation of various behaviors as “play” or “non-play” often seems subjective. One hallmark of play that definitively distinguishes it from other types of
behavior natural to children is the fact that, unlike most behaviors and daily activities, play behavior does not contribute to survival (Lillard, 2015). Generally, children’s play is seen as intrinsically motivating—they play because they enjoy the freedom to choose how and what to play (Ashiabi, 2007). Lillard (2015) elaborates on several comprehensive theories of play to operationalize play as a spontaneous, intrinsically motivated, and ultimately pleasurable or rewarding activity (p. 428). Physical movements during play are often exaggerated, awkward, or clearly distinct from functional or purposeful action. Play behavior can also often be repetitive without requiring a strict form or structure.

Through play, children come to develop a wide variety of skills, reaping social and emotional benefits. Over time, children’s play is observed to be increasingly varied and complex as their social and cognitive skills become more sophisticated. Research has found significant positive correlations between social skills and play quality (Colwell & Lindsey, 2005). Jean Piaget’s theory divides play into three stages of development that parallel children’s cognitive developmental progression: sensorimotor play, symbolic play, and play with rules (Lillard, 2015). The sensorimotor stage lasts from birth to approximately two years of age, and is primarily characterized by a child’s interactions with their environment. During this time, children develop an understanding of their bodies and how they can affect the environment. Children’s earliest play is a pleasurable reaction to a reflexive action, such as sucking their fingers or attending to a shiny object. Soon these accidental impulses become intentional. During the first year, children discover that their behavior has varied effects on the world around them, and they eventually learn to coordinate and integrate different actions to cause desired outcomes in the physical world and on others around them (Gerrig, Zimbardo, Campbell, Cumming, & Wilkes, 2011). The main cognitive goal of sensorimotor play is the development of symbolic
thought, which will eventually enable the child is able to separate an action from its meaning (Lillard, 2015). This advancement is perhaps most evident in the child’s development of object permanence, the understanding that objects exist independent of the child’s awareness, as when during a game of peek-a-boo, the child’s mother is not actually gone although her eyes and other facial features are hiding.

A focus on symbolic play follows the sensorimotor period. The advances made during this stage are striking: while the primary goal of sensorimotor play is to assist the child in understanding her relation to the environment, symbolic play’s impacts are far more broad and can be seen across multiple domains. In the development of symbolic play, the gradual separation of concrete action from meaning allows children’s imagination to be the guiding force in their play. Early on, children use toys in representations of real-world actions, such as using a pretend phone to call Grandma. Later, they will develop the ability to use a wooden block or even their own hand to represent that phone, adding a further layer of abstraction to their play. This developing ability to allow one object or body part to stand in for or symbolize another impacts far more than toy use—while early symbolic play usually takes the form of repetitive rehearsals of real-world events that lack a coherent narrative, children soon develop the ability to create narratives about the progression of events in their play. These symbolic representations in play are, in fact, the foundation for children’s developing abstract thought. As the ability to substitute one object for another or to play a role develops, children engage in increasingly more complicated pretend play incorporating a structured narrative, a number of different characters and a variety of novel situations to be enacted. Children’s ability to narrate or describe their own play is a particularly challenging task that requires the ability to understand and communicate about their own actions (Howes & Matheson, 1992).
Advances in a child’s cognitive development are accompanied by social and emotional developments which also affect the way that children engage with one another in play. As children interact with one another, their emotions are typically highly charged and readily apparent. Over time, as children learn to manage interpersonal relationships, they also learn to communicate and regulate their own emotions. Rather than waiting for a caregiver to respond to their emotional state, a child will begin to seek out assistance, a significant move in the direction of emotional regulation. The ability to self-regulate also plays an important role in children’s social interaction (Ashiabi, 2007). During the preschool years, children’s peer interactions are less closely supervised and structured by adults than they were during infancy and toddlerhood, and it becomes increasingly important that children learn to regulate their own emotions during play and in other situations of social interaction. The negotiation of social interactions is complicated, as it involves the balancing of multiple people’s needs, ideas, and desires. Development in this area is slow but steady, with social interaction becoming increasingly complex over time. Very young children first engage in solitary and onlooker play, in which they focus intently on their own play, or play alone while occasionally stopping to notice what other children are doing. Parallel play is a transition from this individual play to more social play. In parallel play, children play near each other and modify their play based on their observations of their peers. At this stage, they might share toys or occasionally even verbalize to one another. Children’s play then goes through several further stages that are driven by advances in social development, as operationalized by Howes and Matheson (1992). According to the rubric advanced by these theorists, *simple social play* is used to describe play situations where children talk, smile, or exchange toys but do not play together; *complementary and reciprocal play* is operationalized as children’s play in a game involving reciprocal and cooperative roles;
cooperative social pretend play describes situations in which children play pretend roles together; and finally, complex social pretend play describes play which involves cooperative social pretend play and communication between children or narratives made by children about their play.

The growing level of sophistication in play is driven by a child’s social development, but play can also serve as an important vehicle for social learning. Play training has been shown to increase children’s social skills as demonstrated in their interactions with peers (Colwell & Lindsey, 2005). Relationships have also been found between children’s skill at role-playing and their social cognition, meta-cognition, and moral development (Sawyer, 1996). Additionally, children who tend to engage in more sophisticated play have been shown to have social advantages over children whose play is somewhat less complex (Colwell & Lindsey, 2005), and the early development of complex play forms has also been associated with higher rates of prosocial and social behavior and lower rates of aggression and social withdrawal (Howes & Matheson, 1992). In sum, children’s interactions with their peers are key to their development of age-appropriate social skills. Sociodramatic play in particular allows children the chance to practice socioemotional and cognitive skills, including perspective taking, self-regulation, planning behaviors, and problem solving (Asiabi, 2007).

Play is especially important in children’s cognitive and social development as it provides children a safe space to learn about the world—first through repetition and practice and eventually via the creation of novel themes, ideas, and roles. Children’s play is a tool for practicing new and complicated ideas and making sense of the scenarios that they encounter (Von Klitzing, Kelsay, Emde, Robinson, & Schmitz, 2000), but it can also focus on far simpler activities. In early observations at the school profiled in this thesis, a preschool teacher taught a
lesson in which students helped mix pancakes; and for several weeks after this lesson, children continued to pretend to bake pancakes in the classroom. Their play grew increasingly complex over time, to the point where children incorporated pancake cooking into more detailed dramatic play scripts. Play gives children a sense of mastery over their environment. Additionally, through play, children are able to express and process their emotions, and recreate and explore social dynamics (Kyratzis & Ervin-Tripp, 1999; Russ & Fiorelli, 2010). Dramatic play is one of their best tools for comprehending and exploring social rules and expectations.

Dramatic play in particular allows children to explore their notions about the world and experiment with new ideas, including their understanding of their own and others’ identities. In pretend play, “children can be free to express themselves, their ideas, their emotions, and their fantastic visions of themselves, of other people, and of the world,” (Russ, 2014, p. 3).

*Sociodramatic play* is an advanced form of dramatic play and requires highly developed social and cognitive skills. Sawyer (1996) defines sociodramatic play as “pretense play in which [children] enact dramatic roles and fantasy scenarios” (p. 289). However, even in their fantasy play, the themes children explore are strongly related to their own real-world behaviors and their home lives (Von Klitzing et al., 2000). Studies show that while children are aware that their dramatic play is pretend, their pretense does not limit the possibilities for self-reflection during play (Russ & Fiorelli, 2010). In fact, dramatic play allows children the opportunity to create and explore representations of their world in a safe setting (Reunamo et al., 2013).

**Play Narratives**

Children’s involvement in groups and the development of their relationships with peers are both very much fostered by language. Even infants as young as 20 months old work to adjust their interaction styles based on their partner’s responses and actions. Right from the start, there
is a strong connection between children’s language and their relationships (Katz, 2004). Cohen and Uhry (2007) found that while children played together, their speech became more integrated over time—each child started to include phrases and mannerisms in their speech that their partner used. In closely examining the conversations between two pairs of preschool girls, Katz (2004) found remarkably different forms of interaction. While one pair’s dialogue was mostly humorous and nonsensical, the other pair, similar in age, worked to create narratives about their current play and actions. Despite these differences in style between the pairs of girls, both groups had strong bonds and relationships that seemed to be facilitated by their joint narratives. Over time, the individuals involved in each pair began to incorporate words and patterns reflective of their partners’ speech into their own language, which seemed to further facilitate their bonding.

Brown, Roediger, and McDaniel (2014) write that personal narratives allow an individual to step back from a situation and describe and zero in on “what I’m good at, what I care about most, and where I’m headed,” (p. 138). Narratives involve cognitive distance from a situation and the ability to explain that situation as it relates to a broader context. In this way, like schemas, narratives provide a structure for understanding and integrating new experiences into existing knowledge (Brown et al., 2014; Kyratzis & Green, 1997). The operationalization of narrative is complex, as a narrative can involve multiple aspects—people use narratives as a verbal representation of their world and experience (Kyratzis & Green, 1997) and narratives can also be used to create a common script of experience or play. Typically, narratives are used to reflect on experiences, and often they will include background and setting information along with the description of the main event. Narratives are both reflective and generative; they are used to reflect on events and confirm social values, and they also provide a base for the establishment of group identity, values, and norms to take shape (Kyratzis & Green, 1997). The
personal nature of narratives makes them an ideal vehicle to explore and understand individual identity within a larger group (Kyratzis, 1999). Children’s play narratives demonstrate a particularly sophisticated set of skills: a child’s ability to understand and reflect on their own role and actions in play (Howes & Matheson, 1992). The narratives children generate about their own play provide insight into their perspectives, their values, and their worldview beyond what can be understood from their behavior alone.

In order to successfully narrate an event, the narrator must have a solid understanding of the listener’s knowledge of the situation—in other words, they must take the perspective of the listener to understand what information the listener already has and what information the narrative should provide them. For this reason, narration is a skill developed over time: very young children tend to provide much less background information in their narratives than do older children and adults, partly because they are unused to the structure of a narrative, and partly because they might not fully understand that the listener is lacking important background information (Berman, 2001). Children’s narratives develop in complexity over time—older children are significantly more likely to narrate their play than are younger children (Field, De Stefano, & Koewler, 1982). While this age difference is partially driven by differences in language skills, young children also lack the cognitive sophistication to understand how to step back from their play so as to sufficiently describe their actions at that moment or to situate their play within a larger context.

The narratives children generate during play provide an amazingly in-depth view of their perspectives, their values, and their understanding of the world. Play narratives are especially informative and revealing given that play is for many children the primary mode of social activity and exploration. Children narrate their play in several ways. They may act out a role
using a doll, a puppet or other toys as props, or they may enact a role themselves, often planning and negotiating about their behavior, their narrative, with other children (Kyratzis, 1999; Sawyer, 1996). The language of play can include noises and verbalizations to represent different objects or animals, the creation of and naming of objects, and the assignment of roles in play. A variety of markers are used in speech during play to represent different characters—for example, a child might speak in a higher-pitched voice to represent her character, and use her regular voice to discuss the game with her friends—as well as to represent actions and objects (Cohen & Uhry, 2007). Narratives in social situations allow children to position themselves with regards to others in the group and to explore a variety of different personality traits and behaviors (Kyratzis, 1999). Speaking through a made-up character in play is especially useful for self-exploration as this strategy allows children to examine a wide range of roles and identities. Through play, children create an intimacy with the characters that they portray; yet this intimacy is inherently impermanent, allowing children to separate themselves from the responsibilities of their character (Marx & Kyratzis, 1998; Sheldon, 1996).

Children’s narratives are frequently used to negotiate with one another during play, as they create stories, characters, and relations between characters. Children playing together must create their roles and the relationships between roles, and throughout their play they must keep those relationships in mind while at the same time remaining cognizant of their everyday relationships with their playmates (Sawyer, 1996). Additionally, different play contexts require different amounts and levels of sophistication of communication to engage together in a successful play narrative. For example, while some props in a dramatic play area might easily fit into a narrative, another toy such as blocks would be significantly more ambiguous, and
language would be especially important to establish common meaning during play (Cohen & Uhry, 2007).

**Gender Differences in Play: Effects on Children’s Behavior and Narratives**

Given that children frequently recreate their world in play, sociodramatic play often reflects a child’s understanding of cultural values and norms, including gender roles and stereotypes. Maltz and Borker (1983) suggest that as young children self-segregate into friendship groups by sex, groups of boys and girls face different expectations for their play and behavior, and eventually these expectations lead to distinctly different emerging social dynamics in groups of boys and girls. As friendship groups develop, the dynamics encourage increasingly stereotyped behavior (Escalera, 2009). Children’s developing gender stereotypes can be seen when they take on adult social roles in play. Play scenarios demonstrate the kind of adult roles children value, as well as their ideas about the roles that they expect they, too, will one day take on (Kyratzis, 1999). Research has demonstrated a variety of differences in the way that boys and girls play. Overall, boys tend to show stronger preferences for gender-stereotyped activities in their play than do girls (Goble, Martin, Hanish, & Fabes, 2012). Moreover, children with rigid gender schemas have been shown to more frequently choose toys that are appropriately gender stereotyped, and they also tend to show preference for children of their own sex in choosing playmates (Liben & Signorella, 1980; Martin & Halverson, 1981). Coupled with these seemingly gender-based differences is the fact that gender has different levels of salience within classroom contexts: cues for sex roles in certain play areas (e.g., a pretend kitchen or home as compared to an area with blocks and toy fire trucks) tend to lead to gender-stereotypical play, whereas in activities where gender roles are less obvious, such as at a sandbox or water table, play behavior can be very similar between boys and girls (Escalera, 2009). Dramatic play gives children an
opportunity to experiment with various representations of what it means to be male or female in a safe environment (Reunamo et al., 2013). As a result, pretend play is a primary avenue for children to explore different ideas about gender and identity.

In an unfamiliar situation, each of us, young and old, may apply information from a related schema to fill in the gaps in our experience (Martin & Halverson, 1981). For example, a work-related schema may inform a person’s actions on their first day of a new job. In this way, schemas organize past experiences to guide future behavior. Additionally, schemas provide a frame of reference for attention in a new situation; and in many situations, people selectively attend to information that their pre-existing schema indicates will be relevant in the new context (Martin & Halverson, 1981). The rigidity and breadth of a person’s schemas therefore influences how they will process a new situation. Children’s gender schemas typically reflect cultural values, stereotypes, and norms for men and women, but the strength and rigidity of schemas varies between children. Children with rigid gender schemas tend to behave differently in play and other social situations than do children with more flexible schemas. Children with rigid gender schemas more frequently choose toys that are appropriately gender stereotyped, show preference for children of their own sex when choosing playmates, and actually demonstrate better recall for gender stereotypical information than do children with more flexible gender schemas (Liben & Signorella, 1980; Martin & Halverson, 1981). As outlined earlier, boys tend to be more rigid in their gender schemas for males than for females (Smetana, 1986), and this gender difference has been shown to lead to important behavioral differences during play.

Distinctions between girls’ and boys’ play are readily apparent in the research. Martin and Fabes (2001) believe that these gender differences arise, at least in part, because children’s understanding of gender stereotypes and social/societal expectations influence their expectations
of play, and they alter their behavior in play accordingly. Most types of physical play, including rough-and-tumble behavior and large motor activities, are more characteristic of boys’ play than they are of girls’ play (Colwell & Lindsey, 2005). Boys’ play tends to involve more physical contact and aggression, whereas girls’ play is often characterized by calmer, more structured games. Additionally, groups of boys playing together tend to develop a social hierarchy early on, with the order remaining relatively stable over time, while girls’ social hierarchy is more variable over time (Martin & Fabes, 2001). These differences in play style reflect the integration of distinct social expectations for girls and boys into play. Kyratzis (1999) found that characters created and enacted during play suggest children’s own values: girls tend to create loving, gracious, and attractive characters, while boys’ characters are often more physically powerful.

The dominance of gendered stereotypes and expectations in play becomes increasingly evident over time. Martin and Fabes (2001) found that the amount of time that preschool children spent playing with same-sex peers at the start of the school year predicted their level of gender-stereotyped play during follow-up observations in the spring. These same children’s levels of aggressive behavior in play were also found to be related to amount of engagement in same-sex peer interaction earlier in the year—boys’ play with same-sex peers in the fall was positively correlated with aggression in the spring, whereas girls’ play with same-sex peers in the fall was negatively correlated with aggression in the spring. Playing in sex-segregated groups reinforces children’s gender-stereotyped behavior, and the effects of gender-segregated play become compounded over time.

These and many other differences observed between the play of boys and girls are further exemplified in children’s play narratives. Boys’ and girls’ narratives display significant differences in both form and content that reflect stylistic differences in play as well as distinct
group dynamics. Among both boys and girls, narratives are used to confirm membership in friendship groups within the classroom. Exclusivity in friendship groups is often typical in groups of girls. Sheldon (1996) found that some preschool girls use language to exclude others from their group or play. Black (1989) studied the language used between same-sex triads of preschool children and found that girls were more likely to directly address and discuss any conflicting ideas that came up as they structured their play, whereas boys were more likely to suggest new directions or to start playing alone if the group did not respond to their first idea. Escalera (2009) found that while creating extended fantasy narratives, girls frequently engaged in and continued those narratives for long periods, while boys tended to disrupt their own narratives with logistical negotiations. Based on observations like these, Sheldon (1990, in Escalera, 2009) suggested that gender stereotypes push girls’ language to be more cooperative and outwardly oriented, while boys are pushed to express their independence with speech.

Through their form, content, and context, children’s play narratives reflect children’s values and those gender stereotypes that inform their behavior. The narratives preschool children create are typically reflective of peer culture and their own level of identity development as well as the developmental level of the group as a whole. Play narratives highlight children’s understanding of gender and its place in their social world, which includes the classroom environment, their home situation and the larger community and culture (Kyatzis & Green, 1997). In addition to differences in form, the content of play narratives also frequently reflects cultural values related to gender stereotypes. Kyatzis and Green (1997) point to spontaneous narratives of preschool girls as an example of early joint narrative construction. Preschool girls use their narratives to construct a stable group identity, and over time, these narratives come to include or exclude various members and to affirm group membership. Boys’ narratives similarly
reflect gender stereotypical values: while girls’ narratives tend to promote group intimacy, boys often use narratives to negotiate for power and determine social hierarchy (Kyratzis, 1994).

**Gender, Gender Stereotyping, and Play Narratives in a Modern Context**

The present study explored the relations between young children’s formation of and adherence to gender stereotypes, their preferences for specific styles of play, their friendship groups, and the narratives they create during play. Much of the relevant research and theorizing in this area rests on empirical data gathered in the 1980’s and 90’s. Given the significant social changes that have occurred since this time, children’s stereotypes and play behaviors must be re-examined in a modern context. Current social trends have attempted to minimize gender as a significant defining trait of personality, and many preschools have attempted to move away from gendered language, toys and activities (Bigler & Liben, 2007) in an effort to create a more inclusive classroom environment. However, neither the importance of schematic processing as a cognitive tool nor the salience of gender in media and culture can be fully minimized. Despite schools’ best efforts, children continue to be exposed to, if not bombarded by, binary gender stereotypes (Barberá, 2003). The primary goal of this study was to explore the prevalence of and impact of gender stereotyping in a contemporary preschool environment.

Given that all children in the study were above age 2, it was hypothesized that (H1) all study participants would be able to correctly label themselves and others by gender. As gender constancy is an advanced cognitive skill that develops with age, it was also expected that (H2) children’s Gender Constancy Stage would vary by age, with older children being more likely to display higher stages of gender constancy and younger children being more likely to display lower stages of gender constancy. Along with the expected effects of age on gender constancy, it was hypothesized that (H3A) older children would demonstrate stronger gender stereotyping than
would younger children, and that \((H_{3B})\) masculine gender stereotypes in particular would become more rigid with age. Gender was also expected to be a predictive factor for children’s demonstrated levels of gender stereotyping. It was hypothesized that boys would display stronger gender stereotypes than girls. Additionally, regarding specific masculine and feminine stereotypes, it was thought that \((H_{4A})\) boys would display more rigid masculine stereotypes than feminine stereotypes, but that \((H_{4B})\) girls would display more of a balance between masculine and feminine stereotypes. Given that both an understanding of gender constancy and the adherence to gender stereotypes increase with age, it was expected \((H_5)\) that a positive relationship would be demonstrated between children’s Gender Constancy Stage and their adherence to gender stereotypes as operationalized by scores on a Gender Stereotype Test and a Story Stems Test.

It was further hypothesized \((H_{6A})\) that older children’s play would be observed to be more complex than younger children’s play, and that \((H_{6B})\) this effect might be moderated by children’s gender. Furthermore, narrative, a more complex tool used to organize and negotiate play, was expected to be demonstrated \((H_7)\) more often among older children than younger children and among girls than boys, based on previous research demonstrating that the cognitive development levels and verbal sophistication of many preschool girls are more advanced than those of boys. It was also expected \((H_8)\) that all children would play more often with peers of their own sex than with opposite-sex peers, due to the gender group segregation that begins early in life. Finally, it was hypothesized \((H_9)\) that both children’s gender and age would affect their choices for costumes and roles in play. More specifically, it was expected that, in an experimenter-designed play scenario, girls and boys would prefer to wear gender-stereotypical
costumes and act out gender-stereotypical roles, rather than gender-non-stereotypical costumes and roles.

In this investigation, the relations between children’s understanding of gender and their play were further analyzed through close analysis of their play behavior across contexts. Specific hypotheses regarding these observational measures were generated only after the data had been collected, allowing for the construction of composite variables that best captured the behaviors observed during these play sessions. Overall, it was hypothesized that children who in testing were found to adhere more strongly to gender stereotypes would be more likely to play in ways that were stereotypically acceptable for children of their sex. It was further hypothesized that children’s adherence to gender stereotyping would also affect their demonstrated toy, activity, and play partner preferences. Given that boys tend to display more rigid gender schemas than do girls (Smetana, 1986), it was also hypothesized that boys would be more strongly influenced by gender stereotypes in their play. Overall, the expectation was that many of the same gender stereotypes and hallmarks of gender stereotypical play highlighted by earlier studies would again emerge from the data collected at this 2016 state-of-the-art daycare center. However, it was also expected that numerous instances of play that contradict prevailing gender expectations and stereotypes would be observed.

**Method**

**Participants**

The sample consisted of 33 children (18 boys, 15 girls) enrolled in a laboratory preschool on the campus of a private undergraduate college in the Northeast United States. The school was located in a suburban town, and participants were primarily white and from middle to upper class families. As this was a laboratory campus preschool, blanket consent for children to participate
in vetted research studies were given by parents at the start of the school year. Parents did have the option of requesting that their children not take part in this specific study if they so wished.

The school was divided into three classrooms based on children’s age. The older classroom included 16 children who were 4 years old turning 5 (10 males, 6 females; $M = 59.44$ mos., $SD = 3.22$ mos.), and the younger classroom included 17 children who were 3 years old turning 4 (8 males, 9 females; $M = 45.65$ mos., $SD = 3.62$ mos.). The preschool also had a class of 12 two-year-olds who did not participate in the study.

**Setting**

Each classroom was a fairly large space, divided roughly into two separate sections. There were areas for reading, puzzles, art, a water table, a sand box, play dough, manipulative toys (such as small blocks or rubber bands on a pegboard), writing, science exploration, and blocks, as well as an area called the Children’s Museum, where children displayed projects or items they had found outside. The older classroom also had a listening center for audiobooks, a collage table, and a portfolio of each child’s past work.

Each classroom also had an area specifically dedicated to dramatic play. The dramatic play area in the older classroom quite large, with a loft, kitchen setup, table and chairs, etc. The younger classroom had a slightly smaller dramatic play area. Diagrams of the classroom layouts and photographs of the dramatic play areas for each classroom can be found in Appendix B.

Both classes regularly started their day with free play outside. Each class had their own outside area with a climbing structure, slide, swings, digging area, and sand area. The older classroom’s playground also had a functional water pump and tricycles. The younger classroom’s playground had a small indoor greenhouse where children could garden, read books, or draw. After about an hour outside, both classes typically went inside where children had a long block
of time for free play, snack, and a circle meeting, during which time they sang songs and read a story together. At the end of each morning, children went back outside and were picked up by their families on the playground. This schedule was adjusted as needed for weather or special events. Data for this investigation were collected during the winter months. As a result, the daily schedule occasionally involved a much longer block of free play time inside the classroom followed by circle time and snack to end the morning.

**Materials**

A variety of testing materials and observational protocols were employed.

*Gender Constancy Interview*

Given that children’s understanding of gender is closely correlated with their understanding and use of gender stereotypes, (Leinbach, Hort, & Fagot, 1997), the researcher tested children using a slightly modified version of the Gender Constancy Interview (GCT) (Slaby & Frey, 1975). This measure assessed children on their understanding of gender identity, gender stability, and gender consistency across situations. The interview involved 10 items that tested children’s ability to label a person by gender and their ideas about gender stability and consistency over time and across situations. A complete list of these test items can be found in Appendix C. The first four questions on the GCT assessed children’s ability to identify gender based on appearance. Children were presented with four puppets—a boy, a girl, a man, and a woman—one at a time, and were asked to identify the puppet’s gender. In the original GCT, after the child’s had identified a puppet’s gender, a second question further prompted participants by asking if the puppet was the opposite sex of the child’s response—for example, if the child responded that the girl puppet was, in fact, a girl, the researcher would next ask, “Is this a boy?” in reference to the same puppet. This type of follow-up question was eliminated in the current
study due to concerns about potentially confusing children or giving them doubts about their answers. The original GCT also asked children to identify several pictures of people by gender, but these questions were also eliminated due to time constraints. Responses to the first four questions were scored together to measure whether each child had a full understanding of gender labeling—if they answered each of the four items correctly, they received an overall positive score on this dimension.

The test also involved several questions designed to assess children’s understanding of gender stability (questions 5-7). For these questions, children were asked if they were a girl or a boy, whether they were a girl or a boy as a baby, and whether they would be a man or a woman when they grew up. Follow-up questions were then posed to further determine whether children truly understood gender stability. For example, after a girl said that she had been a baby girl, the researcher asked whether she had ever been a baby boy. A full score on the gender stability subscale required that a child correctly identify their own gender in each question, and respond negatively to each follow-up question. Questions 8-10 assessed children’s understanding of gender consistency across situations; for example, boys were asked what gender they would be if they wore girl clothes, and girls were asked what gender they would be if they wore boy clothes. Follow-up questions were not asked for this category. A full score for gender consistency required a correct answer on each of these three questions.

*Story stems*

Children were presented with four story stems. These stories introduced situations in which fictitious children had the potential to act in gender conforming or non-conforming ways. Prior to being presented with each of these four stories, a “practice” story stem featuring a gender neutral character and situation was presented to ensure that children understood the
format of the questioning. Two of the four gendered stories involved children playing with toys that are stereotyped by gender, and two other story stems involved characters behaving in gender-stereotypical ways. The verbal presentation of each story stem was accompanied by a corresponding drawing of a child to represent the story’s main character, and two pictures to represent the potential story outcomes. For example, the practice story stem designed to test children’s comprehension of the task was presented by the researcher using a line drawing of a gender neutral character (an alien). Children were told that the alien could either eat pancakes or bananas for breakfast. Each of these breakfast options was portrayed in a separate line drawing, and the child could point to, pick up, or verbally identify their choice. A complete list of story stems and stimulus puppets and images can be found in Appendix D.

Gender Stereotyping Test

The Gender Stereotyping Test (Leinbach et al., 1997) is a test of children’s adherence to gender stereotypes. The test employed in this investigation was composed of a list of items categorized in pretests as typically masculine, e.g. a snow shovel or a bear, or typically feminine, e.g. a butterfly or a ribbon. Each item was depicted by a simple yet attractive line drawing; a complete set of images can be found in Appendix E. Several list items were abstract characteristics or qualities, but all test items could, in fact, be represented pictorially. For example, angular was represented by a drawing of an angular shape. Because the original Gender Stereotyping Test included more items than could be realistically presented to children in the time allotted for this portion of the study, some items (an equal number of masculine and feminine probes) were eliminated from the protocol. The present research study also added to the protocol several occupations and behaviors that had been categorized as stereotypically masculine and feminine in previous research by Liben and Signorella (1980). A complete list of
items that were included in the testing and the drawings representing each item can be found in Appendix E. The original Gender Stereotyping Test asked children to categorize items as being for boys and men, or for girls and women. This study introduced a third category—items for boys, men, girls, and women—so that children were not forced to categorize items that they had not previously thought of as gender stereotyped.

Observation of free play

Children were observed during free play periods in their classroom and on the playground. Before observation, the researcher created a randomized list of class members and observed the children in that order, without regard to the activity(s) they were engaged in during the observational time period. Each child was observed for four 1-minute intervals—three observation periods inside the classroom and one observation period on the playground. After one minute observing, the researcher recorded information about the child’s play. The child’s activity, sex of their play partners, complexity of social play, and themes observed in play were recorded. As much as possible without disturbing their play, the researcher also recorded examples of the child’s language during observations to look for themes in children’s play narratives.

The coding of these observational data focused specifically on the potential influential effects of gender stereotyping on the content of children’s play. The researcher coded for children’s playmates using the categories employed by Goble et al. (2012). More specifically, Goble and colleagues compared children’s behavior in social contexts to their behavior in solitary play, and in interaction with teachers. Coding categories for children’s play included: solitary play, playing with teachers, playing with male peers, playing with female peers, and playing with both male and female peers (Goble et al., 2012). The researcher also employed a
scale introduced by Howes and Matheson (1992) that drew on the Howes Peer Play Scale 
(Howes & Matheson, 1980) and added additional coding categories to capture the complexity of 
social interactions during play. Previous research has found that children’s preferences for 
gendered activities are strongly influenced by the sex of their play partner; and in this study, 
special attention was paid to this variable. The researcher also coded the sophistication of 
children’s social interaction in play using Howes and Matheson’s scale. If a child did not interact 
with other children during the observation period, their play was coded as solitary play. Parallel 
unaware play was identified in situations when two children played the same activity within 3 
feet of each other but did not interact or acknowledge one another; parallel aware play was 
operationalized as parallel play in which the children made eye contact but did not engage each 
other; simple social play described play situations where children talked, smiled, or exchanged 
toys but did not play together; complementary and reciprocal play was operationalized as 
children’s play in a game involving reciprocal and cooperative roles; cooperative social pretend 
play described situations in which children played pretend roles together; and finally complex 
social pretend play was operationalized as play which involved cooperative social pretend play 
and communication between children or narratives made by children about their play.

Free play observations were also coded for gender stereotyped themes, behaviors, and 
roles that children introduced or enacted during play. In order to compile a list of play features to 
be coded, the researcher spent approximately one half hour in each of the two classrooms 
oberving the children’s play during unstructured free play time. During this time, the researcher 
recorded themes, roles, and behaviors that children enacted in their play, and spoke to each 
teacher about play themes that were typical in the classroom. General categories for play themes 
were then created based on these observations and information provided by teachers. Play coding
categories employed included *domestic themes*, which were further classified as either masculine or feminine based on the children’s actions and roles they created in play; *occupation themes*, again classified as stereotypically masculine or feminine; *animal play*, which included times when a child pretended to be an animal or took care of an animal; and *action/hero* play, when a child acted out a rescuer or superhero role.

*Pre-orchestrated play scenario: Birthday Party Game*

Children were observed playing in assigned same-sex friendship pairs after being introduced to a pre-orchestrated play scenario. Non-controversial pairs of children were created by the preschool director based on observation and teacher input. The researcher provided one king and one queen costume, both equally attractive, and dyadic play was observed in a quiet room separate from the classroom. More specifically, a table was set up with toys and other items children could find at a birthday party. Photos of the costumes and the setup of the birthday party table can be found in Appendix F. Children were told that at this birthday party, kids had been invited to wear costumes. The researcher then explained that two costumes were available, one for a king and one for a queen, and suggested that both of the children could put on a costume and then play birthday party together. These birthday party play sessions were videotaped for later coding.

Play sessions were coded for the effects of gender stereotyping on children’s play. Children’s play tends to reflect their own values and real-world behaviors that they exhibit outside of play (Von Klitzing et al., 2000). Research has found that boys tend to show stronger preferences for gender-stereotyped play activities than do girls (Goble et al., 2012). Children with rigid gender schemas have also been found to more frequently choose toys that are appropriately gender stereotyped over ones that are not, and all children also tend to choose
playmates of their own sex (Liben & Signorella, 1980; Martin & Halverson, 1981). This portion of the research protocol was designed to further explore these connections. Coding involved a close analysis of children’s language, focusing in particular on themes and roles that they introduced in their play, their creation and explanation of their own play characters, and their negotiations and interactions with one another. Specifically, the researcher and a second coder looked for themes such as aggression and caretaking, gendered roles in play such as a mother, princess, superhero, or racecar driver, and gender stereotypical behaviors exhibited in children’s characters.

**Procedure**

The present research study involved several different tests and observation, including a free play observation, several tests about gender stereotypes, and an observation of a pre-fabricated play scenario. The researcher carried out testing over a period of approximately five weeks. The classroom free play observations took place over the first week of testing, but due to time and weather constraints, the outdoor observations took place several weeks later. The next week, each child completed the gender stereotype test session. Upon completing that phase of testing, the researcher was able to begin observations with the pre-fabricated play scenario at the end of the second week, and those observations were completed within the third week.

*Free Play Observation*

Each child was observed four times for 1-minute segments, for a total of 4 minutes over a period of several days. During each of these observations, information about a child’s activities, level of social interaction, and sex of the peer(s) with whom they were interacting was recorded as detailed previously. Teacher presence and interaction was also noted. This procedure combined observation protocols from Howes and Matheson (1992) and Goble et al. (2012).
Gender Stereotype Testing

Children’s comprehension of gender stereotypes was tested using the Gender Constancy Interview (Slaby & Frey, 1975), story stem completion, and the Gender Stereotype Test (Leinbach et al., 1997). A full transcript of the dialogue and instructions employed during this testing session can be found in Appendix G. Throughout the entire testing process, in addition to the experimenter, a second adult was present to serve as a coder and to keep track of children’s responses. All testing in this session was audio recorded in order to facilitate accurate coding at a later date. The researcher invited children one at a time to leave their classroom and come with her to a quiet room where a table was set up with all necessary props and testing materials. The researcher sat the child comfortably at the table, and told the child that she had a game to play that she thought the child would really enjoy.

First, the researcher tested each child using the Gender Constancy Interview (GCI) developed by Slaby and Frey (1975). The original GCI involved the presentation of a boy and girl doll, and a man and woman doll. In the present investigation, the researcher used basic male and female puppets (a simple line drawing of a child, laminated). The researcher asked the child to identify one puppet at a time as a boy or girl, or a man or a woman. The test continued with questions about the child’s understanding of gender across time and situations. A complete list of questions can be found in Appendix C.

Next, the researcher explained that she was going to tell the child the beginning of a story, and that the child could decide what happened next in the story. Four story introductions involved situations where the characters could act in gender conforming or nonconforming ways. Children were first presented with a “practice” story featuring a gender neutral character and neutral situation to ensure that the child understood the game. The researcher read each story
stem, accompanying the reading with a picture of a boy or girl to represent the child character in the story (or an alien to represent the gender neutral character in the example), and illustrations of the choices that could be made. After hearing each story, the child was asked to tell or show the researcher what happens next, using their own words and the provided drawings. An example story stem would be, “This is Johnny (show boy puppet). Johnny goes to his classroom in the morning. He sees that there is a spot to play with a truck, like this (show picture of toy truck), and a spot to play with a baby doll, like this (show picture of baby doll). Tell or show me what happens next.” A full script of story stems and illustrations can be found in Appendix D.

Finally, children were tested individually on their adherence to gender stereotypes using the Gender Stereotype Test (Leinbach et al., 1997). Children were asked to identify items in a list of behaviors, traits, and actions as being masculine, feminine, or neutral. The researcher explained to the child, “I am going to show you pictures of things people could look at, or play with, or use to work with. Some of these things are more for girls and women, some of them are more for boys and men, and some of them can be for both boys and girls. If you think the picture is mostly for girls, it goes in this box. If you think the picture is mostly for boys, it goes in this box. If you think the picture is mostly for everyone, it goes in this box. You can tell me which box to put the cards in or you can put the cards in yourself.” Each basket was accompanied by an image—a woman and two girls for the women/girls box, for example—so that children could identify which basket was for each category. These images can be found in Appendix E. After hearing these instructions, children were presented with several example items (a man, a woman, and a hand) to make sure they understand the task. It was presumed that every child would sort the woman into the girls/women box, the man into the boys/men box, and the hand into the neutral box. If a child was confused on this section, the researcher re-explained the rules, and
asked the child again to sort the sample pictures. After going through the sample items, the researcher moved on to the actual test items.

The researcher showed each child one item at a time and identified verbally each image for the child. For the first several test items, the researcher told the child, “If you think the picture is mostly for girls, it goes in this box. If you think the picture is mostly for boys, it goes in this box. If you think the picture is mostly for everyone, it goes in this box.” Once the child seemed to understand the game, the researcher stopped explaining which box was for which category. If the child hesitated to respond to an item, the researcher reminded the child that they could put the card in the box themselves, or they could tell the researcher which box to put it in or simply point to a box. The researcher verbally announced children’s responses for purposes of coding. After the test was completed, the researcher thanked the child for playing her game and offered them a choice of two stickers. The researcher then walked the child back to his or her classroom.

*Pre-Orchestrated Play Scenario: Birthday Party Game*

Children were observed playing in same-sex/same-age friendship pairs that were created by the preschool director to be non-adverse. One same-sex pair of children at a time was brought by the primary investigator to a separate, quiet room at the preschool that had been preconfigured to facilitate play and video recording. A second researcher was waiting in the room with a video camera to record the play session. The room was set up with a table with toys and other items children might find at a birthday party, as well as a king and a queen costume. The researcher told children that they could each wear a costume and play birthday party together. She also explained that she was interested in the kinds of things that kids do when they play, and that her friend would be making a video of them playing.
If one or both children objected to wearing a costume, the researcher encouraged costume use by saying that this was a game with costumes, but also made clear that it was their choice to wear the costume or not. The researcher then told the two children that she had work to do while they played, and that she thought they would have fun with the game. If children appeared to have a hard time starting to play, the researcher prompted their play by asking them who might go to a birthday party, and what they might do there. She then sat quietly in a chair and took note of any themes or characters that emerged in the play. After about 7 minutes, or when the children were seemingly done with the game (whichever came first), the researcher thanked the children and explained that they should leave the costumes and the props where they had found them so that others could also have a chance to play. She then let each child choose one of two stickers and took them both back to their classroom. A full transcript of the session dialogue/protocol can be found in Appendix H.

**Coding**

*Gender Stereotype Test*

The Gender Stereotype Test consisted of 30 items, half of which were stereotypically masculine and half of which were stereotypically feminine. Each child was presented with the items one at a time and asked to sort the individual items as belonging more with men and boys, belonging more with women and girls, or neutral (belonging with men, boys, women, and girls). If a child sorted an individual item according to prevailing gender stereotypes, that was counted as a correct response for that item. Sorting an item as neutral or in opposition to common gender stereotypes—for example, sorting the color pink as belonging with men and boys—was counted as an incorrect answer.


**Story Stems Task**

The Story Stems Task presented children with four stories, two with a boy main character and two with a girl main character. In each story, the character was presented with a choice between a masculine option and a feminine option. The participant was asked what they thought the character would choose. If the child answered according to prevailing gender stereotypes, their answer for that story was counted as correct. If the child answered with the opposite option, neither option, or both options, their response was counted as incorrect.

**Classroom Observations**

Brief notes were taken during each 1-minute observation period to preserve important details in the child’s speech or action; and at the end of each observation interval, the researcher filled out a coding sheet which allowed for a more complete recording of details from the observation. A sample observation coding sheet can be found in Appendix I. The researcher noted how many boys and girls the child in question had played with, whether the child had been interacting closely with a teacher, and which activity the child was engaged in. To facilitate coding, two lists of potential activities had been constructed beforehand, one to suit indoor observations and one focused on outdoor activities, based on the stations that were included in the classroom and the playground set-up. Indoor activities included: reading; play-dough; art; blocks; dramatic play; science table; water table; and sand table. Outdoor activities included: slide or climbing structure; swings; sandbox (a category which included both sandboxes and digging in dirt); running/large motor play; and dramatic play. While there was no dramatic play area specified on the playground, the researcher remained close enough to the children to note whether their play involved dramatic play elements, such as fabricated characters or a running narrative. Activity coding also allowed for an ‘Other’ category for activity. In the instance that a
child’s play did not match a specific pre-defined category, the researcher noted the child’s specific activity. Because a large portion of those “other” activities involved a child playing with Legos or other manipulative toys, during data analysis the decision was made to create a separate coding category that would capture this type of fine-motor play. During outdoor observations, if the child’s play was coded as “running/large motor activities”, the researcher noted what specific activity the child was engaged with. Because outside observations took place during the winter months, this running/large motor activities category often included interactions with ice or snow on the playground—for instance, children attempting to slide on ice, or pretending to ice skate. During data analysis, a separate coding category was created for snow play due to the prevalence of this activity.

Coding also involved more specific details regarding the children’s play. The researcher first recorded the child’s overall level of interaction with their playmates that had been observed during play, using a list of coding categories borrowed from the work of Goble et al. (2012) and Howes and Matheson (1992). These categories are explained in depth in the Materials section. **Complex social pretend play** signified pretend play in which children narrate or explicitly discuss the narrative of their play. For example, a child might announce to his playmate that he is pretending to be a dog. In **cooperative social pretend play**, children engage in pretend play without explicitly discussing the narrative of their game. A child pretending to be a dog might demonstrate that idea by barking and crawling on the ground, but would not explain their actions. Like the other two categories, **complementary reciprocal play** involves a significant level of interaction in play, and children engage with one another to create a coherent game. However, this style is marked by realistic play that does not involve fantasy or pretend elements. Throughout the observation and coding process, the researcher paid close attention to children’s
mannerisms and language in order to distinguish between these three similar but distinctive types of play.

Finally, the researcher coded for any themes that were evident in play. Based on general observations of each classroom, a list of general thematic categories was created. The researcher coded for domestic play, occupation play, animal play, and action/superhero play. Domestic play and occupation play were both subdivided into masculine and feminine categories. Themes were coded based on children’s explicit play narrations and their general speech directed to one another in play, their actions, and the props that they used. Often, children did not explain their play verbally. In these cases, themes were coded based on overt actions. For example, play that involved a child pushing a digger truck and making construction noises was coded as a masculine occupation theme, and the researcher noted that construction was specifically included. If children primarily engaged with toys and did not evidence any themes in their play, the researcher noted this as well. In data analysis, these demonstrations of individual themes were recoded into new overarching categories in an effort to capture a broader range of behavior. These new thematic categories encompassed construction (including children’s discussion of building and explicit demonstrations through action of building—such as a child creating a tower in the blocks area); doctor/medical care, which came up frequently due to the presence of medical-themed toys in each classroom’s dramatic play area during observation; food, including cooking and eating, primarily based around kitchen set-ups in the dramatic play areas; animals, which could involve a child pretending to be an animal or pretending to take care of one; transportation, including driving, construction vehicles (when the focus of play was on moving the vehicles rather than on construction per se), trains, and airplanes. Several different types of play collectively fell under the umbrellas of caring or aggressive themes in play. Domestic
masculine roles were rarely observed in play, while feminine domestic roles such as cooking or taking care of a baby were frequent. *Caring* behavior included any caretaking and domestic activities such as cooking food. Instances in which a child pretended to be a medical professional were also coded as caring behavior because typically, children’s play focused on domestic caring/healing interactions rather than medical issues or instruments. *Aggressive* themes were coded when children were either verbally or physically aggressive with one another, or when they tried to control another child’s play through the imposition of verbal instructions.

*Pre-Orchestrated Play Scenario*

Each session of the Birthday Party Game was video recorded to allow for detailed analysis. Videotapes were coded by the primary researcher and a second trained coder. The researcher and the second coder each coded video segments individually and then compared their data. Any disparities in coding were discussed until agreement was reached, and therefore there was no need to compute levels of inter-rater reliability. This coding of the tapes focused on the play session as a whole, the introduction of the game and costumes, and several select intervals during the actual play period. Because two children were involved in each session, the researcher and coder completed the coding of one child’s behavior before moving on to coding for the second child. A complete display of coding categories and levels for this game may be found in Appendix A.

The researcher and the assistant coder watched the child as they were introduced to the game by the researcher, beginning from the time when they first entered the testing room. During this introductory phase, two costumes were presented, a king costume and a queen costume, and the child’s response to the costumes was coded on several levels. More specifically, the coders recorded whether each child wore a gender-stereotypical or non-stereotypical costume and crown
(coded separately), or whether the child chose not to wear either the costume or the crown. Some children chose to wear the Queen costume backwards—while the costume was a dress, it opened in the back, and wearing it backwards made it look more like a robe or a cape. If the child wore the queen costume, a variable to indicate whether the costume was backwards or forwards was coded for as well. The child’s enthusiasm for each costume was reverse coded on a 1-5 point scale, from 1 (very enthusiastic) to 5 (very unenthusiastic), with a score of 3 indicating a general neutral response to the costume. This assessment was based on the child’s verbal response as well as their affect. Another variable addressed whether or not the child had attempted to negotiate for a specific costume, and if they had, whether their reasoning included gender-focused thinking (e.g., “I’m a girl, so I want the queen costume,”) or gender-neutral reasoning (e.g., “I love purple, so want the purple costume,”). If, as part of this reasoning process, the child said either “king” or “queen” (or, sometimes, prince/princess), their negotiation was coded as gender-focused. Throughout the remainder of the video, the researcher and second coder paid attention to whether each child played a character, and where character portrayal was involved, they also included in their coding whether the role was gender stereotypical, non-stereotypical, or neutral, and whether the role matched the child’s costume.

Rather than attempt to code each entire play session, three 20-second intervals were sampled—the first and last 20 seconds of play, as well as a 20-second interval in the middle of the play session. For each of these three intervals, several variables were coded. Researchers addressed the child’s affect (rated on a 1-5 scale, from extremely positive to extremely negative); whether their play was realistic, fantasy-based, or a mix; if they held more, less, or equal power in their interactions with their play partner—or whether they did not interact sufficiently to demonstrate power balance; and their narration level (high, medium, or low). A coding of low...
narration indicated either that the child did not speak at all during the play session or spoke very few words. Narration was counted as medium when a child responded to questions or statements from their peer or the researcher, but did not initiate conversation. A coding of high narration indicated that the child spoke a great deal throughout the play session and that this use of language was specifically directed at communicating the play’s script/direction to the play partner. Ratings of affect were based both on the child’s language and on their demonstrated enthusiasm for the game as inferred in terms of participation level, vocalizations, and body language.

Results

Quantitative Data

Gender Constancy

Children typically first show a firm understanding of gender identity/labeling, and only later come to understand gender stability, and finally gender consistency. Mastery of all three of these areas of understanding indicates that they fully comprehend the principle of gender constancy. One of the hypotheses driving the present study ($H_1$) was that all children, regardless of age group or gender, would be able to accurately label themselves and others by gender. As predicted, each child interviewed correctly identified him or herself as a boy or a girl. This finding was to be expected, given that all children in the study were above the age of 2, the developmental milestone indicated in the literature when most typically developing children have mastered this understanding. Additionally, 29 out of 33 children answered correctly each of the four gender identity questions. Three of the four children who made errors mislabeled only one photo. The fourth child in this group answered each labeling question incorrectly, with the
experimenter noting that based on his tone and mannerisms, he appeared to know the correct answers but was purposely giving incorrect responses.

Table 1 displays a Guttman scale for the Gender Constancy Scores earned by children in this study. This table shows the number of participants at each stage of the developmental progression underlying the understanding of gender constancy. The atypical patterns displayed by five children are also depicted. This group’s scores are categorized as Non-Stage and four distinct response patterns are noted. Given this variety of responses, it did not make sense to include these five cases in subsequent data analyses that incorporated Gender Constancy Stage as an independent / grouping variable.

<table>
<thead>
<tr>
<th>Type</th>
<th>Gender Identity</th>
<th>Gender Stability</th>
<th>Gender Consistency</th>
<th>Number of Children</th>
<th>Age (By Classroom)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Boys</td>
<td>Girls</td>
<td>Both</td>
<td>Younger</td>
<td>Older</td>
</tr>
<tr>
<td>Stage:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>+</td>
<td>—</td>
<td>—</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>+</td>
<td>+</td>
<td>—</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Non-Stage:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>—</td>
<td>—</td>
<td>+</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>—</td>
<td>+</td>
<td>+</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>+</td>
<td>—</td>
<td>+</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>—</td>
<td>+</td>
<td>—</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

For the purpose of further analysis, the number of children at each Gender Constancy Stage was computed as a proportion out of 28 (the total number of children displaying typical stage-wise patterns of gender constancy development) and these proportions were compared using a Chi-square Goodness of Fit Test. Results showed that there was no significant difference.
in the number of children performing at each of the three stages of gender constancy understanding, \(X^2(2, N = 28) = 2.64, p = .267\). However, an effect size estimate (\(\phi = .500\)) indicated that within this small sample, there was, in fact, a meaningful difference in the number of children scoring at each of the three stages. As shown below in Figure 1, the largest number of children scored at Gender Constancy Stage 3, the highest of the three stages of gender constancy understanding, and the lowest number of children scored at Gender Constancy Stage 1, the lowest stage of understanding.

![Figure 1. Frequency of children at each Gender Constancy Stage.](image)

It was expected (H2) that older children would be more likely to score at high Gender Constancy Stages and that younger children would be more likely to score at low Gender Constancy Stages. An independent samples t-test was conducted to compare Gender Constancy Stage in older and younger children. Age groupings were determined by children’s classroom placement. Children in the younger class were 3 and 4 years old (\(M = 45.65\) mos., \(SD = 3.62\) mos.), and children in the older class were 4 and 5 years old (\(M = 59.44\) mos., \(SD = 3.22\) mos.). There was a significant difference in Gender Constancy Stage found between children in the younger classroom (\(M = 1.93, SD = .917\)) and children in the older classroom (\(M = 2.57, SD = 1.97\)).
SD = .51); \( t (20.43) = -2.29, p = .033, d = 0.86 \). As a Levene’s test indicated unequal variances (\( F = 8.01, p = .009 \)), a separate variance estimate t-value was computed and degrees of freedom were adjusted from 26 to 20.43. Older children were significantly more likely to be at higher stages of gender constancy comprehension, while younger children were more likely to be at lower stages of gender constancy comprehension (see Figure 2).

![Figure 2](image)

*Figure 2.* Frequency of children from the younger and older class at each Gender Constancy Stage.

In order to look for gender differences between children at different Stages of Gender Constancy, an independent samples t-test was conducted to compare boys and girls on the highest stage of gender constancy attained. There was no significant difference in Gender Constancy Stage between boys (\( M = 2.29, SD = .73 \)) and girls (\( M = 2.21, SD = .89 \)); \( t (26) = .23, p = .818 \). An effect size estimate, Cohen’s-\( d \), calculated at .09 further supported this conclusion. Figure 3 illustrates the number of boys and girls who presented at each Gender Constancy Stage.
Knowledge of Gender Stereotypes

Children’s adherence to gender stereotypes was measured by their responses to the Gender Stereotype test and the Story Stems task. Individual items on each measure addressed either masculine or feminine stereotypes. Collapsing across the two measures, each child received three scores regarding their stereotype adherence: adherence to masculine stereotypes, adherence to feminine stereotypes, and adherence to gender stereotypes overall. For the purpose of clarity, children’s score for adherence to masculine stereotypes will be referred to as their Masculine Stereotype Score (MSS), their score on adherence to feminine stereotypes will be referred to as their Feminine Stereotype Score (FSS), and their adherence to all gender stereotypes—both masculine and feminine—will be termed their Composite Stereotype Score (CSS).

Each response to an item on the Gender Stereotype Test was coded as correct if the child had correctly sorted the item according to prevailing gender stereotypes; items on the Story Stem Task were similarly coded as correct if the child chose the story ending that followed traditional gender stereotypes (see Methods for further information). Each child’s Masculine Stereotype
Score was calculated as a percentage of their correct responses to masculine stereotyped items on the Gender Stereotype Test and to the two masculine story stems. Each child’s Feminine Stereotype Score was calculated as a percentage of their correct responses to feminine stereotyped items on the Gender Stereotype Test and to the two feminine story stems. Children’s Composite Stereotype Scores were then calculated as an average of their Masculine Stereotype Score and their Feminine Stereotype Score.

A third major hypothesis driving this investigation ($H_{3A}$) was that older children would show evidence of more rigid gender stereotyping in their responses than would younger children, and in particular that ($H_{3B}$) older children would display higher levels of masculine stereotypes than would younger children. As shown in Table 2 below, a marginally significant ($t(31) = 1.31$, $p = .20$) age difference was found for children’s Composite Stereotype Score, with older children scoring higher ($M = .47$, $SD = .15$) than younger children ($M = .39$, $SD = .19$). The Cohen’s-$d$ effect size estimate for this between-group comparison was .47 and met Cohen’s (1992) convention for a moderate effect ($d = .50$). In other words, age had a noticeable and meaningful effect on children’s adherence to gender stereotypes. A marginally significant age difference ($t(31) = -1.68$, $p = .10$) was also found for adherence to masculine stereotypes (see Table 2). Older children ($M = .57$, $SD = .23$) had higher scores than younger children ($M = .44$, $SD = .23$) on the Masculine Stereotype measure. The Cohen’s-$d$ effect size estimate for this between-group comparison was .55 and met Cohen’s (1992) convention for a moderate effect size ($d = .500$). The effect of age on adherence to feminine stereotypes ($t(31) = -.32$, $p = .76$, $d = .149$) was not significant. Taken together, these findings indicated that this increase in adherence to gender stereotypes over time was primarily the result of an age-driven tendency to subscribe more and more strictly to masculine stereotypes.
It was also predicted that children’s levels of gender stereotyping would be driven, at least in part, by their gender. More specifically, it was hypothesized that (H4A) boys would display more rigid masculine stereotypes than feminine stereotypes. Girls, on the other hand, were expected (H4B) to show more balanced stereotyping across the two genders, yielding similar Masculine and Feminine Stereotyping Scores. As shown in Table 3, there was no significant gender difference \((t(31) = -0.05, p = .603)\) found between boys \((M = .42, SD = .15)\) and girls \((M = .45, SD = .20)\) on their Composite Stereotype Score. However, a marginally significant gender difference \((t(31) = 1.62, p = .116)\) was found for adherence to masculine stereotypes, with boys \((M = .56, SD = .22)\) scoring higher than girls \((M = .43, SD = .25)\). The Cohen’s-\(d\) effect size estimate for this between-group comparison was .55 and met Cohen’s (1992) convention for a moderate effect. Additionally, there was found a highly significant \((t(31) = -3.22, p < .005, d = 1.13)\) gender difference for adherence to feminine stereotypes, with girls \((M = .47, SD = .22)\) scoring higher than boys \((M = .27, SD = .12)\).
It was also hypothesized that (H₃) Gender Constancy Stage would be systematically related to levels of gender stereotyping, and it was also tentatively hypothesized that this relation would be specific to age or gender groupings. To test these predictions, children’s Gender Stereotype Scores were compared to their Gender Constancy Stage, illustrated below in Figure 4. The proportion of correctly answered masculine stereotype items to correctly answered feminine stereotype items was computed separately for children scoring at each of the three Gender Constancy Stages. Z-scores were then computed to test for significant differences between these proportions. For children at Stage 1, the Z-score was .30, \( p = .764 \). For children at Stage 2, the Z-score was .64, \( p = .522 \). For children at Stage 3, the Z-score was .59, \( p = .555 \). Although these tests failed to reveal any significant differences between proportions, it did appear that children were more likely to adhere to masculine stereotypes than feminine stereotypes. Collapsing across age and gender categories, this conclusion was, in fact, strongly supported. A paired t-test revealed that children’s Masculine Stereotypes Scores (MSS) \( (M = .50, SD = .24) \) were significantly higher than their Feminine Stereotypes Scores (FSS) \( (M = .36, SD = .20) \), \( t (32) = 2.99, p = .005 \). The Cohen’s \( d \) for this calculation was .63.

<table>
<thead>
<tr>
<th>Type</th>
<th>Girls</th>
<th></th>
<th>Boys</th>
<th></th>
<th>( t(31) )</th>
<th>( p )</th>
<th>Cohen’s ( d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite Stereotype Score</td>
<td>.45</td>
<td>.20</td>
<td>.42</td>
<td>.15</td>
<td>-.53</td>
<td>.603</td>
<td>0.17</td>
</tr>
<tr>
<td>Masculine Stereotype Score</td>
<td>.43</td>
<td>.25</td>
<td>.56</td>
<td>.22</td>
<td>1.62</td>
<td>.116</td>
<td>0.55</td>
</tr>
<tr>
<td>Feminine Stereotype Score</td>
<td>.47</td>
<td>.22</td>
<td>.27</td>
<td>.12</td>
<td>-3.22</td>
<td>.003</td>
<td>1.13</td>
</tr>
</tbody>
</table>
Figure 4. Percentage of masculine and feminine stereotypes correct from children at each Gender Constancy Stage. Error bars represent standard error of the mean.

In an effort to further understand these results, the medians for Masculine, Feminine, and Composite Stereotype Scores were calculated and used to split children into high- or low-stereotyping groups for each of these three categories. Using chi-square-based tests, the associations between children’s standing in these high-low groups and their understanding of gender constancy were then explored (see Table 4). A Chi-square Test of Independence crossing Gender Constancy Stage with high and low Feminine Stereotyping Score categories showed no significant association between these two variables, $X^2 (2, N = 28) = 2.57, p = .276$. However, an effect size estimate for this test ($r = 0.30$) indicated a moderate effect size. Gender Constancy Stage also failed to show a significant association with children’s scores on Masculine Stereotypes, $X^2 (2, N = 28) = 2.33, p = .312$. However, once again, an effect size estimate for this test ($r = 0.29$) indicated a moderate effect. Finally, children’s Composite Stereotype Scores were crossed with Gender Constancy Stage. These results approached statistical significance,
\[ X^2 (2, N = 28) = 5.22, p = .074. \] An effect size estimate for this calculation \((r = 0.43)\) indicated a moderate effect.

Table 4

<table>
<thead>
<tr>
<th>Gender Constancy Stage</th>
<th>Stereotype Type</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Masculine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 1</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Stage 2</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Stage 3</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Non-Stage*</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

*While non-stage children are included in this chart, they were excluded from the related analyses.

Qualitative Data: Effects of Gender Stereotypes on Play

It was hypothesized \((H_{6A})\) that older children’s play would be more complex than younger children’s play. It was also expected \((H_{6B})\) that gender might moderate this effect. A Chi-square Test of Independence that crossed Complexity of Peer Interactions in Play (CPIP) observed during naturalistic observation with age categories showed no significant association between these two variables, \[ X^2 (2, N = 132) = -.06, p = .970. \] An effect size estimate for this test \((r = 0.02)\) also did not indicate an important association. Although it would be expected that complexity of play would be tied to age, this association was not found with these data. Figure 5 displays the frequencies for the number of children in the Oldest and Middle classrooms within each category of play complexity.
Complexity of Peer Interactions in Play was next compared between genders. A Chi-square Test of Independence crossing CPIP during naturalistic observation with gender categories demonstrated a marginally significant association, $X^2 (2, N = 132) = 5.79, p = .055$. An effect size estimate for this test ($r = .21$) indicated that there was a small but meaningful association between gender and CPIP. Figure 6 displays the frequencies of the number of boys and girls observed to engage in each of the three levels of peer interaction in play (tallied across all four naturalistic observations).

Figure 5. Complexity of peer interactions in play during observations by age.
It was hypothesized ($H_7$) that older children would display more narratives in their play than would younger children, and that girls would display more play narratives than would boys. As part of the coding of children’s behavior during the Birthday Party Game, each child received three scores for their narration level displayed at three different checkpoints. Narration level was reverse coded—a score of 1 signified high use of narration (as operationalized in Methods), a score of 2 signified moderate use of narration, and a score of 3 signified low use of narration. Each child’s three individual narration scores were added together to produce an overall Narration Score. There was a marginally significant age difference found for narration levels ($t(30) = 1.39, p = .18$), with older children scoring slightly lower ($M = 6.81, SD = 1.60$) than younger children ($M = 7.63, SD = 1.71$). The Cohen’s-$d$ effect size estimate for this between-group comparison was .50, and met Cohen’s (1992) convention for a moderate effect. Keeping in mind that these data were reverse coded, age had a noticeable and meaningful effect on children’s narration levels, with older children narrating more than younger children. There was also found a significant difference between boys and girls for narration levels ($t(30) = -2.15, p < .05, d = .740$), with girls ($M = 7.86, SD = 0.95$) scoring higher than boys ($M = 6.72, SD = 1.71$).

![Complexity of Peer Interactions in Play by Gender](image)

**Figure 6.** Complexity of peer interactions in play by gender.
SD = 1.96). Again, because these data were reverse coded, these test results indicate that boys narrated their play more often than did girls.

It was hypothesized (H₈) that all children would play more often with peers of their own sex than with peers of the opposite-sex. Paired-samples t-tests were conducted to compare the total number of girls to the total number of boys that each child played with over the course of all four naturalistic observations. Among girls, there was not a significant difference (t(14) = .50, p = .625) between number of female playmates (M = 3.20, SD = 2.11) and number of male playmates (M = 2.87, SD = 1.88); the Cohen’s-d for this comparison was 0.17, which approached a small effect size. Among boys, there was also not a significant difference (t(17) = -1.32, p = .204) between number of female playmates (M = 2.00, SD = 1.68) and number of male playmates (M = 2.72, SD = 1.49); however, the Cohen’s-d for this comparison was 0.45, which approached a moderate effect size—boys were more likely to play with boys than with girls. The difference in effect sizes for these tests focused on girls and boys indicates that boys were more likely to prefer playmates of the same gender, while girls showed less of a preference based on gender. It was further expected that this tendency to play with peers of the same gender would be stronger for older than for younger children (as operationalized by child’s classroom placement). Results from a series of paired (dependent) t-tests are displayed below in Table 5. A significant difference was found only for younger boys (t(7) = -2.76, p = .028, d = 0.964), indicating that this group was significantly more likely to play with boys than with girls.
Table 5

Gender of Children’s Playmates By Child’s Age and Gender

<table>
<thead>
<tr>
<th>Age/Gender</th>
<th>Female playmates</th>
<th>Male playmates</th>
<th>df</th>
<th>t</th>
<th>p</th>
<th>Cohen’s-d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Younger girls</td>
<td>2.22</td>
<td>1.86</td>
<td>1.89</td>
<td>1.45</td>
<td>8</td>
<td>.385 .710</td>
</tr>
<tr>
<td>Younger boys</td>
<td>1.50</td>
<td>0.93</td>
<td>2.75</td>
<td>1.58</td>
<td>7</td>
<td>-2.76 .028*</td>
</tr>
<tr>
<td>Older girls</td>
<td>4.67</td>
<td>1.63</td>
<td>4.33</td>
<td>1.51</td>
<td>5</td>
<td>.291 .783</td>
</tr>
<tr>
<td>Older boys</td>
<td>2.40</td>
<td>2.07</td>
<td>2.70</td>
<td>1.49</td>
<td>9</td>
<td>-.326 .752</td>
</tr>
</tbody>
</table>

*Indicates a significant p-value

It was hypothesized (H₀) that both children’s gender and age would affect their choices for costumes and roles in play. More specifically, it was expected that girls would be more likely to prefer the queen costume, whereas boys would prefer the king costume. Given that children played the birthday game in pairs, they were frequently unable to choose and wear a same-gender stereotypical costume. To circumvent this experimental design artifact, costume preference was analyzed in terms of children’s demonstrated enthusiasm for each costume, rather than in terms of their actual costume choice. Levels of enthusiasm for the costumes were coded on a scale from 1 to 5, with 1 signifying the least enthusiastic response and 5 signifying the most enthusiastic response. As shown in Table 6 below, a comparison of the enthusiasm shown by boys and girls for the queen costume showed a highly significant difference (t (30) = -4.02, p < .001), with girls (M = 4.07, SD = .92) showing significantly more enthusiasm than boys (M = 2.61, SD = 1.09). The Cohen’s-d effect size estimate for this between-group comparison was 1.45, and met Cohen’s (1992) convention for a very large effect. Additionally, a comparison of boys’ and girls’ level of enthusiasm for the king costume showed a highly significant difference (t (30) = 2.07, p < .05), with boys (M = 3.89, SD = 1.23) demonstrating significantly
more enthusiasm than girls ($M = 3.07, SD = .92$). The Cohen’s-$d$ effect size estimate for this between-group comparison was 0.76, and met Cohen’s (1992) convention for a large effect.

Table 6

<table>
<thead>
<tr>
<th>Costume</th>
<th>Girls</th>
<th>Boys</th>
<th>$t(30)$</th>
<th>$p$</th>
<th>Cohen’s $d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Queen</td>
<td>4.07</td>
<td>2.61</td>
<td>-4.018</td>
<td>&lt;.001</td>
<td>1.45</td>
</tr>
<tr>
<td>King</td>
<td>3.07</td>
<td>3.89</td>
<td>2.074</td>
<td>.047</td>
<td>0.76</td>
</tr>
</tbody>
</table>

As demonstrated in Table 7, there were no significant age differences for enthusiasm for the queen costume ($t(30) = 1.00, p = 1.000$) or for the king costume ($t(30) = 1.39, p = .175$). However, the Cohen’s-$d$ effect size estimate for a between-group comparison of older and younger children’s enthusiasm for the king costume was 0.49, which approached Cohen’s (1992) convention for a moderate effect.

Table 7

<table>
<thead>
<tr>
<th>Costume</th>
<th>Younger</th>
<th>Older</th>
<th>$t(30)$</th>
<th>$p$</th>
<th>Cohen’s $d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Queen</td>
<td>3.25</td>
<td>3.25</td>
<td>1.00</td>
<td>1.000</td>
<td>.00</td>
</tr>
<tr>
<td>King</td>
<td>3.81</td>
<td>3.25</td>
<td>1.39</td>
<td>.175</td>
<td>.49</td>
</tr>
</tbody>
</table>

It was also expected that boys would be more likely to take on a gender-stereotypical role than a neutral or non-stereotypical role in their play, while girls were expected to display less of
a strong preference for gendered roles. Twenty-two children took on a pretend role/portrayed a pretend character during the Birthday Party Game. For this analysis, the pretend play roles adopted by children were coded as either gender stereotypical or neutral/non-stereotypical. A chi-square analysis crossing children’s gender with the type of roles they chose to take on during play showed no significant association between these two variables, \( \chi^2 (1, N = 22) = .182, p = .670, r = .091 \). Children’s preference for gender stereotypical or non-stereotypical roles was also crossed, in a separate chi-square analysis, with age. This test also failed to demonstrate any significant association between these two variables, \( \chi^2 (1, N = 22) = .210, p = .647, r = .098 \).

Two composite scores (MBCS and FBCS) designed to capture children’s gender-stereotyped behavior across all naturalistic play observations as well as the Birthday Party scenario were calculated. As a first step, each individual measure that contributed to these composite scores was recoded as a dichotomous variable (present/absent) with presence signifying that the child had been coded as engaging in the measure-targeted behavior. For example, a boy would be given a score of “present” if he engaged in aggressive behavior during any one of the four naturalistic observations. The Masculine Behavior Composite Score (MBCS) included: whether the child played with boys during naturalistic observations; presence of transportation themes in their play during naturalistic observations; presence of construction themes in their play during naturalistic observations; presence of aggressive themes/behavior in their play during naturalistic observations; whether they played with blocks during naturalistic observations; whether they played with Legos during naturalistic observations; whether they were enthusiastic about the king costume during the Birthday Party Game; whether they attempted to negotiate for a costume based on gender during the Birthday Party Game; and whether they played a masculine character during the Birthday Party Game. The Feminine
Behavior Composite Score (FBCS) included: whether the child played with girls during naturalistic observations; presence of animal themes in their play during naturalistic observations; presence of food themes in their play during naturalistic observations; presence of caring themes/behavior in their play during naturalistic observations; whether they did art during naturalistic observations; whether they read books during naturalistic observations; whether they were enthusiastic about the queen costume during the Birthday Party Game; whether they attempted to negotiate for a costume based on gender during the Birthday Party Game; and whether they played a feminine character during the Birthday Party Game. These MBCS and the FBCS composites each yielded a total possible score of 9.

It was expected that each child’s MBCS and FBCS Scores would be systematically and negatively related to one another. A paired samples t-test was conducted to explore this possibility, comparing scores on the MBCS ($M = 3.88, SD = 2.20$) and FBCS ($M = 3.61, SD = 1.85$). Contrary to expectation, no significant difference was found between the children’s two scores ($t(32) = .44, p = .666, d = .13$). Because this difference between the MBCS and the FBCS might have been be determined at least in part by the child’s gender, girls’ and boys’ scores were also analyzed separately via additional paired samples t-tests. A significant difference was found between girls’ scores for MBCS ($M = 2.33, SD = 1.40$) and FBCS ($M = 4.40, SD = 2.10$), $t(14) = 3.27, p < .005, d = 1.160$. A paired samples t-test conducted to compare boys’ MBCS ($M = 5.17, SD = 1.92$) and FBCS ($M = 2.94, SD = 1.35$) scores also showed a significant difference, $t(17) = -2.72, p = .016, d = 1.344$. These analyses revealed that throughout naturalistic observations and the pre-fabricated play scenario, girls earned higher Feminine Behavior Composite Scores than Masculine Behavior Composite Scores, and that boys earned higher Masculine Behavior Composite Scores than Feminine Behavior Composite Scores.
Furthermore, the differences in effect size estimates showed that the magnitude of the difference between expressed gender stereotypical and non-stereotypical behavior in play was larger for boys than it was for girls—all children were more likely to express gender stereotypical behavior than gender non-stereotypical behavior, but the difference was more obvious in boys’ play.

It was expected that children’s MBCS and FBCS composite scores would be significantly and positively related to their Masculine and Feminine Stereotype Scores. Because it was also thought that these relations would be gender-specific, boys and girls were analyzed separately. A Pearson product-moment correlation coefficient was computed to assess the relation between MBCS and Masculine Stereotype Scores. Among boys, there was no significant correlation between the two variables, $r = -.05, n = 18, p = .830$. Among girls, there was also no significant correlation between the two variables, $r = .26, n = 15, p = .356$. A Pearson product-moment correlation coefficient was also computed to assess the relation between FBCS and Feminine Stereotype Scores. Once again, there was no significant correlation between the two variables found for girls ($r (15) = .19, p = .497$) or boys ($r (18) = .06, p = .826$).

It was also expected that children’s MBCS and FBCS composite scores would be significantly and positively related to their Gender Constancy Stage. Omitting study participants who had shown atypical patterns of Gender Constancy Stage development, children were categorized as displaying high levels of Gender Stereotypical Behavior if they scored at or above the median composite score for their gender group (MBCS or FBCS). Children scoring below the median were categorized as displaying low levels of gender stereotypical behavior. Using the same median split technique, children were also categorized as displaying high or low levels of Gender Non-Stereotypical Behavior (i.e. behavior stereotypically associated with the opposite
gender). A chi-square analysis crossing the three levels of Gender Constancy Stage with these high-low GSB categories failed to show a significant association between these two variables, $\chi^2 (2, N = 28) = 2.174, p = .337$. However, this comparison demonstrated a moderate effect size ($r = .279$). As demonstrated in Figure 7 below, children at Gender Constancy Stage 3 were more likely to display high levels of Gender Stereotypical Behavior than low levels of Gender Stereotypical Behavior. This trend was less apparent among children at Stage 2, and there appeared to be a reverse effect at Stage 1. A similar chi-square analysis comparing Gender Constancy Stage to children’s high-low gender non-stereotypical behavior also failed to demonstrate any significant association between these two variables, $\chi^2 (2, N = 28) = .952$, $p = .621$. However, this comparison demonstrated a small, but meaningful, effect size ($r = .184$). As demonstrated in Figure 7 below, children at all Gender Constancy Stages were more likely to display high levels of Gender Stereotypical Behavior than to display high levels of Gender Non-Stereotypical Behavior. Children at all Gender Constancy Stages were also less likely to display high levels of Gender Non-Stereotypical Behavior than to display low levels of Gender Non-Stereotypical Behavior. Taken together, these observations mean that children were more likely to act in gender stereotypical than non-stereotypical ways.

![Figure 7](image-url)

*Figure 7.* Gender stereotypical and non-stereotypical behavior compared to Gender Constancy Stage.
Discussion

Children’s Understanding of Gender

Children’s understanding of gender was initially assessed through the Gender Constancy measure. Beginning around age four, children begin to understand gender constancy, the idea that gender is a stable identity trait and does not change based on context or superficial markers (Lippa, 2002). The findings in this study support past research showing that children are able to correctly identify themselves by gender by the age of two. Additionally, the majority of children in this study were able to correctly identify others by gender. Importantly, however, gender labeling and identity questions only partially address the concepts involved in gender constancy; also important are the concepts of gender stability (understanding that one’s own identity is stable) and gender consistency (understanding that gender is a stable identity regardless of outward appearance or activities). An understanding of these latter concepts involves far more sophisticated levels of cognition and tends to show a sequential progression, with stability coming before consistency. Most (85%) of the children who participated in this study displayed a typical developmental pattern of gender constancy understanding. More children scored at Stage 3 than at Stage 1 of Gender Constancy; and there was found a significant age difference, with older children more likely to score at Stage 3 and younger children more likely to score at Stage 1. Boys and girls did not tend to score differently on the Gender Constancy Test. A comparison of these data with past findings reported by Slaby and Frey (1975), the creators of the original Gender Constancy Scale utilized here, showed strikingly similar patterns. Slaby and Frey (1975) tested children whose ages ranged from 35 to 67 months—essentially the same range studied here; and across the two studies, an approximately equal percentage of children fell into each Gender Constancy Stage (see Figure 8).
In the present study, children’s knowledge of gender stereotypes was assessed through the Gender Stereotype Test and the Story Stems task. Scores from the two measures were combined to give each child an overall score for adherence to masculine stereotypes, adherence to feminine stereotypes, and adherence to gender stereotypes in general. Gender is made salient to children every day, and over time, they develop schemas to organize and make sense of the information they are gathering about this construct. A great many previous investigations have revealed that the majority of children show an increasing rigidity in their gender schemas over time, with children’s attitudes and behaviors being increasingly dictated over time by their ideas as to what is and is not appropriate for boys and girls. While, in this investigation, an independent t-test failed to reveal a significant difference between older and younger children for Composite Stereotype Score, a Cohen’s-$d$ calculation of a moderate effect size ($d = .50$) showed that age had a noticeable and meaningful impact on children’s adherence to gender stereotypes overall. In fact, this pattern of Null Hypothesis Statistical Procedures failing to reveal significant between-group differences followed by the calculation of effect size estimates that indicated the
presence of meaningful effects was frequently repeated across the data analysis process. At issue here was the small \( (n = 33) \) sample size and the fact that while the significance levels of hypothesis tests are very much driven by sample size, effect size estimates are largely unaffected by the fact that a small, medium or large sample was tested.

In addition to examining children’s overall Composite Stereotype Scores, age differences were also explored separately for masculine and feminine stereotypes. A marginally significant age difference \( (p = .10) \) and a moderate effect size estimate \( (d = .55) \) showed that older children were more likely than younger children to adhere strongly to masculine stereotypes. The effect of age on adherence to feminine stereotypes was not significant and yielded an especially low effect size. Given that age differences were found in relation to children’s adherence to masculine stereotypes but not feminine stereotypes, it might be concluded that as their understanding of the social world develops, children only become increasingly rigid in their ideas about what is appropriate or inappropriate for males. This finding has potentially important real-world significance. As children come to understand social expectations and learn to monitor their own and others’ behavior based on their stereotypical ideas about masculinity and femininity (Martin & Halverson, 1981), it would appear that girls may enjoy far more options than will boys as to how they might look, dress or behave. In fact, Scott (1984) suggests that given current social standards and expectations for males and females, girls stand to gain far more than do boys from transgressing gender stereotypes.

While proportions of children scoring at the various Gender Constancy Stages remained fairly stable from 1975 to 2016, published data reported by a variety of researchers over the decades indicated that adherence to gender stereotypes did not necessarily display the same stability. For this reason, children’s adherence to masculine and feminine stereotypes in this
investigation was compared to findings from Leinbach, Hort, and Fagot’s (1997) original study of adherence to gender stereotypes. Figure 9 presents data on adherence to masculine and feminine stereotypes in 1997 and 2016, coded separately for girls and boys as well as collapsed across gender categories. Both boys and girls showed decreased levels of masculine and feminine stereotypes from 1997 to 2016. An overall negative trend in adherence to masculine and feminine stereotypes over the years was also evident when data from boys and girls were combined.

Figure 9. Adherence to gender stereotypes in 1997 and 2016 by gender.

As outlined earlier, the original Gender Stereotype Test was modified for the present study to include a gender neutral response category so that children would not be forced to give a gendered label to items that they had previously thought of as neutral. However, this modification in testing protocol meant that a direct comparison of the current data to previously reported findings for this measure was impossible. In Leinbach et al.’s original study (1997), children were presented with two options on the Gender Stereotype Test. In other words, if they
had guessed on every item, there was a 50% chance that their response to any particular question would be coded as correct. In contrast, the present study gave children three options for answers on the Gender Stereotype Test. For these children, guessing on every item would give them a 33.33% chance that their response to any particular question would be coded as correct. In order to compare directly the two data sets, for both studies level of adherence to stereotypes was re-calculated as the percentage difference from chance; those differences are displayed in Figure 10. Z-scores were then calculated to compare the proportion of adherence to stereotypes in 1997 to adherence to stereotypes in 2016 for both masculine and feminine stereotypes. These calculations did not reveal a significant change in boys’ or girls’ adherence to masculine stereotypes over the years. However, boys demonstrated a significant decrease in adherence to feminine stereotypes, $Z = 2.15, p = .03$. Girls did not demonstrate a similarly significant decrease in adherence to feminine stereotypes, although given the small sample of girls ($n = 15$), it is important to note that a marginally significant difference was found, $Z = 1.26, p = .208$. When boys’ and girls’ data were combined, there was no significant difference found between the adherence to masculine stereotypes shown by children in 2016 and 1997, $Z = 1.22, p = .222$. There did emerge, however, a highly significant difference between the 1997 and 2016 samples for adherence to feminine stereotypes, $Z$-score $= 2.57, p = .010$. From 1997 to 2016, preschoolers’ adherence to masculine stereotypes showed somewhat of a decrease and their adherence to feminine stereotypes decreased significantly.
The preschool from which participants were sampled for this study serves a fairly homogenous population. The majority of children in the study were white and upper middle class, and many of their parents were professors or other professionals with advanced degrees. Owing to its association with a liberal arts college, the preschool itself is a unique environment in that its curriculum and policies are frequently evaluated and modified to connect with current psychological and educational research. Children often participate in research studies, and are surrounded by a high number of teachers and college student teacher assistants. The preschool strives to create as gender-neutral an environment as possible—teachers avoid gendered language, and encourage children to participate in activities regardless of gender. In previous research, when gender was made salient in the classroom, children tended to adhere more strongly to gender stereotypes in their beliefs and behavior (Hilliard & Liben, 2010). Following this logic, one would expect that in an environment that de-emphasizes gender such as the

*Figure 10.* Difference from chance: Adherence to gender stereotypes in 1997 and 2016 by gender.
preschool studied here, children would be less rigid in their gender schemas as compared to children in a typical preschool setting. Given this laboratory preschool’s efforts to remain gender neutral, coupled with the shifts in cultural norms surrounding gender that have occurred over the past 20 years, it might be expected that the number of children observed to hold strict ideas about gender would be diminished in the current sample. In fact, it appears that children’s adherence to masculine and feminine stereotypes has decreased since Leinbach et al. originally conducted similar research in 1997. This pattern is evident in both girls’ and boys’ adherence to stereotypes. However, although it has been 40 years since the publication of Slaby and Frey’s findings (1975), the percentages of children scoring at each Gender Constancy Stage at this progressive laboratory preschool in 2016 matched almost exactly the percentages reported in the original research. Rigid gender roles and rules may not be emphasized to children in the same ways that they once were, yet it appears that children continue to develop strong ideas about gender differences.

The perpetuation of gender constancy across time and historical context might be explained in terms of multiple factors. First, preschool is largely a protected and controlled environment. Children’s speech and behavior is highly monitored at school, and their toys, books, and surroundings are all carefully structured for them. However, preschoolers obviously do not spend all of their time at school. They go out into the world with their families, where they are exposed to a variety of different cultural values. Even at home, siblings or parents might react strongly to a child who transgresses gender stereotypes. Messages about the broader world are further delivered via television, movies, and books that a preschooler might access at home. For example, a current children’s TV show features a team of dogs who work together in emergency situations. The cast only includes one girl dog and it is this character’s job to fly a
helicopter. Never does she engage with the other male dogs as they complete their rescue missions. Even in a fantastic world where dogs can operate machinery and communicate with one another and with humans, girls are relegated to less active and less dangerous work than are boys (Duran & Whitney, 2015). These gendered messages from books and media are pervasive.

But there is a second important reason why even children learning in gender-neutral classroom settings would be expected to continue to form fairly binary gender schemas. Simply stated, prevailing American norms and expectation continue to be extremely gender-driven. However, in order to successfully come to understand and navigate the social world as it is, children must learn to delineate between and female categories. They must continue to form somewhat rigid gender stereotypes as well as a solid understanding of gender constancy until such a time (should it ever happen) that society becomes truly gender neutral. In fact, Lippa (2002) argues that if we were to create a completely gender-neutral society, children might not develop these same stereotypes and schemas.

Importantly, however, data analyses presented here indicate that in spite of society’s continued emphasis on gender differences, preschoolers living and learning in 2016 may well be forming and subscribing to fewer gender stereotypes than did their same-age peers 20 years ago. Although children were found to be developing an understanding of gender constancy at a rate commensurate with that reported in decades past, it appeared that their development of gender stereotypes and schemas, most especially feminine stereotypes, might not be as rigid or restrictive as they have been reported to be in the past. This conclusion is tempered by the fact that the format of the primary measure used in the 1997 study was not the same as the format employed in this investigation. As described earlier, statistical efforts were made to circumvent this problem and, while conversions to percentages and calculations of the degree to which
children were responding at levels above chance were somewhat cumbersome, the findings surrounding these comparisons across cohorts remain compelling. Importantly, like the preschoolers who participated in this 2016 investigation, Leinbach and colleagues (1997) also investigated a sample of children attending a college campus laboratory preschool; but in this earlier investigation data were also collected from a second perhaps more heterogeneous daycare setting.

Given the development of schemas over time, as well as children’s growing understanding of gender constancy, it was expected that children’s Gender Constancy Stage might be related to their adherence to gender stereotypes. While there was no significant difference found at any specific Gender Constancy Stage between children’s adherence to masculine or feminine stereotypes, the data did reveal an overall trend of children subscribing more rigidly to masculine than feminine stereotypes. This apparent pattern was investigated further by collapsing across age, gender, and Gender Constancy Stage to compare children’s scores for adherence to masculine and feminine stereotypes. A paired t-test revealed a significant effect ($p = .005$) and moderate effect size ($d = .63$), with children adhering significantly more to masculine than feminine stereotypes. This finding held true across Gender Constancy Stages, and was driven primarily by the responses of older children and boys.

Overall, children in this study were more likely to adhere strongly to masculine stereotypes than to feminine stereotypes. The effects of schemas and stereotypes on behavior are powerful, impacting not only playmate and toy preferences (Liben & Signorella, 1980; Martin & Halverson, 1981) but also activity level, demonstrations of aggression, and the establishment and maintenance of social hierarchies (Martin & Fabes, 2001). While two children might begin preschool with very similar interests, preferred activities, and play styles, their behavior may
evolve to be significantly different based on their gender. In this study, the findings showed that children demonstrated stronger stereotypes—and therefore more rigid schemas—for boys than for girls. In other words, at least in this preschool setting, boys’ play may in fact have been more significantly impacted by gender schemas than was girls’ play. While it is generally assumed that girls must conform to especially rigid gender stereotypes and expectations, data from this investigation indicate that for very young children, boys might actually face the brunt of gender stereotyping.

**Evidence of Gender Schemas in Play**

Children’s play was first examined in terms of quantitative measures designed to capture the complexity of children’s interaction with their peers, as well as their narration levels. Children’s play is very much driven by their cognitive ability and tends to become more complex over time. Complexity of play is especially tied to language development, as language allows children to communicate and negotiate with one another regarding play content, roles, and narratives. In the present study, the expected association between age and complexity of peer interactions in play was not demonstrated. However, narration levels were tied to age, with older children observed to narrate their play more than younger children. It is interesting that while overall complexity of play was not associated with age in the current investigation, children’s narration in the Birthday Party Game was associated with age. This disparity may, in fact, stem from a flaw with the study’s design. In the classroom and on the playground, where naturalistic observations were completed and complexity of peer interactions was coded, children were able to choose their play areas, and could decide whether to interact with peers or play alone. The classroom and the playground provided many opportunities for solitary play. In contrast, in the Birthday Party Game, children were invited into the game with a peer. While they could choose
how much they wished to speak during play, they were necessarily placed in a situation where communication and narration were almost inevitable. The fact that older children were more likely to demonstrate high levels of narration than younger children in the Birthday Party Game indicates that the complexity of play within the context of the Birthday Party Game scenario was higher among older children than among younger children.

Measures of complexity of peer interactions (demonstrated during naturalistic observation) and level of narration (measured during the Birthday Party Game) were also compared between girls and boys. A marginally significant gender difference was found, with boys displaying higher levels of simple interactions in their play, and girls demonstrating more instances of complex interactions. Interestingly, during the Birthday Party Game, boys were observed to have higher levels of narration than girls. It was expected that girls’ peer interactions would be more complex than boys’, and this gender difference was demonstrated across naturalistic observations. However, the fact that boys demonstrated higher levels of narrative was not expected. Boys might have been less shy in the new environment of the Birthday Party Game than girls, and therefore more likely to talk. Narrative content was unfortunately not addressed in the present study. Narrative content is highly reflective of narrative complexity—talking about roles and themes in play is more complicated than simply explaining actions. Future research should address narrative content in order to gain a more full understanding of children’s narratives in dramatic play as well as possible gender differences in this area.

Social learning theory states that children’s tendency to act in gender stereotypical ways results from their imitation of same-sex adults in their lives, particularly those whom they view as especially powerful or nurturing (Lippa, 2002). Gender differences in children’s behavior therefore arise in large part from the gendered behavior that children see modeled. In most
situations, when children act in ways that align with prevailing gender stereotypes, they are rewarded; and, in this way, gendered behavior is reinforced and ingrained over time (Lippa, 2002). For this reason, it was expected that older children would display more gender-stereotypical roles and behavior in their play than would younger children. Also, given the overall tendency to adhere more strongly to masculine than feminine stereotypes, it was expected that boys would be more affected (or restricted) in their play by gender stereotypical expectations than would girls. In fact, these hypotheses were supported in that children at Stage 2 and 3 of Gender Constancy were more likely to display high than low levels of gender-stereotypical behavior. Additionally, children at every Gender Constancy Stage were more likely to display low than high levels of gender-non-stereotypical behavior.

Overall, data from the present study reveal a number of important ways in which children’s gender and their adherence to gender stereotypes impacted their actual behavior. A between-group comparison of Composite Stereotype Scores showed no significant difference between girls and boys. However, analyses focused on specific gendered behaviors exhibited during both naturalistic play observations and an experimenter-created play scenario demonstrated that boys displayed stronger adherence to masculine stereotypes than did girls, and girls displayed stronger adherence to feminine stereotypes than did boys. The tendency to develop stronger beliefs about in-group identities is well documented throughout the literature and is replicated in the present study. The themes that children demonstrate in dramatic play have been shown to be reflective of their real-world behaviors and their home lives (Von Klitzing et al., 2000). Through play, children can create and explore representations of their real world in a safe environment (Reunamo et al., 2013). Based upon the findings from the present study, it is possible to conclude that children were able to represent their beliefs about gender in
their play. While children in this investigation tended to adhere less strongly to gender stereotypes than did children in past years, they continued to demonstrate developmentally typical understandings of gender constancy. It is possible that despite the best efforts of this progressive preschool in particular and modern culture overall to limit children’s exposure to gender stereotypes, preschoolers cannot be fully exempt from gendered expectations.

Methodological Limitations

The findings and conclusions presented here should be understood in the context of a fairly atypical preschool environment. Preschool participants’ families were fairly homogenous—most children were white, came from upper-middle class background, and tended to have parents who had earned advanced degrees. The preschool itself is unique in that it is a laboratory preschool located on a college campus. The staff is intimately familiar with cutting-edge research and pedagogical innovations, and regularly incorporate this understanding into their curriculum planning and their teaching style. Class sizes are small, and there is an unusually high teacher-to-student ratio. As explained previously, an overarching goal at the preschool is to create a classroom environment that is as gender neutral as possible. Gendered language from teachers is discouraged, children are never separated by gender for play purposes or other activities, and toys are not labeled as being ‘for boys’ or ‘for girls’. The preschool teachers attempt to make every child feel welcome and accepted regardless of gender, play preferences, or appearance. These unique program characteristics, coupled with the ramifications of a small sample size \(n = 33\), make the data nearly impossible to generalize to a larger population. Additionally, the older classroom was comprised of 10 boys and 6 girls. Because of this gender imbalance, any comparisons between the two classrooms might be biased in favor of boys’ scores.
While children were individually tested on their knowledge of gender stereotypes and gender constancy, the Birthday Party Game scenario, yielding some of the most important behavioral measures in the study, necessitated interaction with a peer. Play partners for the game were selected by the preschool director based on classroom observations and information from teachers. Every effort was made to match children with a partner with whom they regularly interacted and felt comfortable. However, this testing was done in a separate room and it is likely that children’s level of comfort with the game and with their play partner impacted their behavior in the game. The addition of a play partner further complicates the interpretation of the data in that children’s narration levels were likely impacted by an chatty or a silent play partner. Moreover, because children were given a choice of only two costumes to be worn during the game, it was inevitable that many study participants would end up unhappy with their costume. While it was absolutely necessary in this study to make gender salient by providing two especially gendered costumes, this forced choice situation meant that one play partner was often more vocal about the costume they desired, and the other less assertive child in the pair may well have been made uncomfortable by wearing a costume about which they were unenthusiastic.

**Future Directions**

A first step for future research in this area would be to emphasize the incorporation of a more diverse sample. While it proved interesting and somewhat informative to compare contemporary and archival data, it is impossible to know which of the many sociocultural changes that have occurred over the past 20 or 40 years contributed significantly to the differences observed in this investigation. A replication of this study with a far broader, more diverse sample of children would make it possible to generalize findings across a variety of preschool environments and demographic groups. In recent years, a small but growing number of
preschools have attempted to create a more gender-neutral classroom environment. A replication of this study across a variety of preschool environments characterized by differing levels of gender salience, would do much to increase our understanding and awareness of the effects of gender stereotypes on children’s play behaviors, choice of playmates, and emerging sense of self.
References


### Appendix A

<table>
<thead>
<tr>
<th>Play Type</th>
<th>Realistic (focus on toys)</th>
<th>Mixed</th>
<th>Non-realistic (focus on fantastical elements)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Queen</td>
<td>Enthusiastic</td>
<td>Neutral</td>
<td>Non-enthusiastic</td>
</tr>
<tr>
<td>King</td>
<td>Enthusiastic</td>
<td>Neutral</td>
<td>Non-enthusiastic</td>
</tr>
</tbody>
</table>

#### Operationalization

<table>
<thead>
<tr>
<th>Scale Possibilities</th>
<th>Queen Enthusiastic</th>
<th>King Enthusiastic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly</td>
<td>Toy</td>
<td>Toy</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>Toy</td>
<td>Toy</td>
</tr>
<tr>
<td>Neutral</td>
<td>Toy</td>
<td>Toy</td>
</tr>
<tr>
<td>Unenthusiastic</td>
<td>Toy</td>
<td>Toy</td>
</tr>
<tr>
<td>Strongly</td>
<td>Role</td>
<td>Role</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>Role</td>
<td>Role</td>
</tr>
<tr>
<td>Neutral</td>
<td>Role</td>
<td>Role</td>
</tr>
<tr>
<td>Unenthusiastic</td>
<td>Role</td>
<td>Role</td>
</tr>
</tbody>
</table>

#### Coding for Birthday Party Game

- Did the child argue for one costume or the other?
- Which costume was the child expressing a role/character for?
- Did the child express their gender/role?
- Were negotiability raised on the costume?
- Did the child initiate conversation or merely respond?
- Was the role a gender stereotype?
- Did the child play a role?
- Was the role gender stereotypical for the child's gender?
- Did the child play with toys?
- Did the child extend play beyond toy play?
- Was the role gender stereotypical for the child's gender?
Appendix B
Classroom Diagrams

Younger Class

Dramatic Play Area

Block Area
Playground
Older Class

Dramatic Play Area
Classroom Setup
Block Area

Playground
Appendix C
Gender Constancy Interview

Test Items
1—2. (For a boy puppet and for a girl puppet):
   Is this a girl or a boy?
3—4. (For a man puppet and a woman puppet):
   Is this a woman or a man?
5. Are you a girl or a boy?
   Are you a [opposite sex of subject’s first response]?
6. When you were a little baby, were you a little girl or a little boy?
   Were you ever a little [opposite sex of subject’s first response]?
7. When you grow up, will you be a mommy or a daddy?
   Could you ever be a [opposite sex of subject’s first response]?
8. If you were [opposite sex of subject, i.e. “boys” or “girls”] clothes, would you be a boy or a girl?
9. If you played [opposite sex of subject] games, would you be a girl or a boy?
10. Could you be a [opposite sex of subject] if you wanted to be?

Puppet Stimuli

2 http://www.necord.org/boys-coloring-pages/little-boy-coloring-pages/
3 http://www.resimkoy.xyz/ae6ec5d81bd0882e.html
4 http://www.resimkoy.xyz/c98419e613a09190.html
Appendix D
Story Stems Test

**Story Stems**

- Neutral story: This is Wuba *(show alien puppet)*. Wuba goes to the kitchen in the morning. Wuba sees that there are pancakes to eat *(show picture of pancakes)*, and there are bananas to eat *(show picture of bananas)*. Tell or show me what happens next.
- This is Danny *(show boy puppet)*. Danny goes to his classroom in the morning. He sees that there is a spot to play with a toy truck, like this *(show picture of toy truck)*, and a spot to play with a baby doll, like this *(show picture of baby doll)*. Tell or show me what happens next.
- This is Alice *(show girl puppet)*. Alice goes to her classroom in the morning. Alice sees that there is a spot to play with a toy bunny, like this *(show picture of toy bunny)*, and a spot to play with a toy hammer *(show picture of toy hammer)*. Tell or show me what happens next.
- This is Scott *(show boy puppet)*. Scott plays with his family. Scott can help cook dinner *(show picture of food)*, or Scott can play outside *(show picture of a ball)*. Tell or show me what happens next.
- This is Melissa *(show girl puppet)*. Melissa plays with her dog. Melissa can give her dog a treat *(show picture of dog with treat)*, or Melissa can play a game with her dog *(show picture of a dog with a ball)*. Tell or show me what happens next.

---

**Boy, Girl, and Alien Puppets**

![Boy, Girl, and Alien Puppets](https://thegrandnarrative.files.wordpress.com/2011/05/stick-figure-family.jpg)

![Alien](http://www.clker.com/cliparts/K/d/4/o/t/g/alien-outline-hi.png)
### Image Stimuli

<table>
<thead>
<tr>
<th>Banana</th>
<th>Pancakes</th>
<th>Toy truck</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baby doll</td>
<td>Toy bunny</td>
<td>Toy hammer</td>
</tr>
<tr>
<td>Cooking dinner</td>
<td>Playing with a ball</td>
<td>Give dog a treat</td>
</tr>
<tr>
<td>Playing fetch with dog</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

> http://cliparts.co/search?q=hug&btnicon.x=0&btnicon.y=0
## Appendix E

### Gender Stereotyping Test

<table>
<thead>
<tr>
<th>Test Items</th>
<th>Masculine Items</th>
<th>Feminine Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Items from original Gender Stereotyping Test (Leinbach et al., 1997)</td>
<td>Hammer Firefighter hat Bat and ball Snow shovel Bear Angular Grasshopper Airplane Truck Blocks Blue</td>
<td>Needle and thread Broom Baby bottle Heart Pink Flowers Butterfly Feather Curved Ribbon Rabbit</td>
</tr>
<tr>
<td>Incorporated items from Gender Stereotypes and Attitude Scale (Liben &amp; Signorella, 2012)</td>
<td>Rake leaves Fly a plane Fix a car Race a car</td>
<td>Bake cupcakes Buy clothes Buy groceries Clean the house</td>
</tr>
</tbody>
</table>

### Masculine Items

- Hammer
- Firefighter hat
- Bat and ball
- Snow shovel
- Bear
- Angular
- Grasshopper
- Airplane
- Truck
- Blocks
- Blue

---

<table>
<thead>
<tr>
<th>Airplane</th>
<th>Truck</th>
<th>Blocks</th>
<th>Blue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rake leaves</td>
<td>Fly a plane</td>
<td>Fix a car</td>
<td></td>
</tr>
</tbody>
</table>

**Feminine Items**

<table>
<thead>
<tr>
<th>Needle and thread</th>
<th>Broom</th>
<th>Bottle</th>
<th>Heart</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flowers</td>
<td>Butterfly</td>
<td>Clean the house</td>
<td>Feather</td>
</tr>
<tr>
<td>Curved</td>
<td>Ribbon</td>
<td>Rabbit</td>
<td>Pink</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>Bake cupcakes</td>
<td>Buy clothes</td>
<td>Buy groceries</td>
<td></td>
</tr>
</tbody>
</table>

**Images to label boxes for sorting**

- **Things that are for girls and women**
- **Things that are for boys and men**
- **Things that are for everyone**
### Appendix F

**Photographs of Play Scenario**

<table>
<thead>
<tr>
<th>Costume</th>
<th>King</th>
<th>Queen</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Costume</strong></td>
<td><img src="http://www.amazon.com/Forum-Novelties-Crown-Purple-Costume/dp/B002PANEDI?ie=UTF8&amp;psc=1&amp;redirect=true&amp;ref_=od_aui_detailpages00" alt="King Costume" /></td>
<td><img src="http://www.amazon.com/Rubies-Red-Heart-Princess-Costume/dp/B004RQMD80?ie=UTF8&amp;psc=1&amp;redirect=true&amp;ref_=od_aui_detailpages03" alt="Queen Costume" /></td>
</tr>
<tr>
<td><strong>Crown</strong></td>
<td><img src="http://www.amazon.com/Forum-Novelties-Regal-Crown-One-Size/dp/B002YDGR3A" alt="King Crown" /></td>
<td><img src="http://www.amazon.com/Forum-Novelties-Ruby-Heart-Tiara/dp/B003FQEKM0?ie=UTF8&amp;psc=1&amp;redirect=true&amp;ref_=od_aui_detailpages01" alt="Queen Crown" /></td>
</tr>
</tbody>
</table>

9. Only robe was included in game
10. Only dress was included in game
Birthday Party Table
Appendix G
Gender Stereotype Testing Transcript
The experimenter brought the child to a space outside the classroom and seated him/her at a table. A second adult, serving as a coder, was also present in the room.

Experimenter: I’m glad you are going to play a game with me, [child’s name]! My name is Riley, and this is my friend [coder’s name]. [Coder’s name] is going to do some work in this room while we play a game together. During this game, I’ll have a recorder so that I remember what we talk about.

Gender Constancy Interview

Experimenter: In this game, I’m going to ask you some questions.

Each of the first two questions listed was repeated for a male and female puppet, for which illustrations can be found in Appendix D. The researcher put the puppet on the table in front of the child before posing each accompanying question. Once the child answered, the researcher removed the puppet and put the next puppet down.

1—2. (For a boy puppet and for a girl puppet):
   
   Is this a girl or a boy?

3—4. (For a man puppet and a woman puppet):
   
   Is this a woman or a man?

The rest of the questions were directed to the child in the order they are presented here. After the child appeared to be done answering a question, the researcher moved on to the next question.

5. Are you a girl or a boy?
   
   Are you a [opposite sex of subject’s first response]?

6. When you were a little baby, were you a little girl or a little boy?
   
   Were you ever a little [opposite sex of subject’s first response]?
7. When you grow up, will you be a mommy or a daddy?

Could you ever be a [opposite sex of subject’s first response]?

8. If you were [opposite sex of subject, i.e. “boys” or “girls”] clothes, would you be a [opposite sex of subject’s first response]?

If you wore [opposite sex of subject] clothes, would you be a [opposite sex of subject’s first response]?

9. If you played [opposite sex of subject] games, would you be a girl or a boy?

If you played [opposite sex of subject] games, would you be a [opposite sex of subject’s first response]?

10. Could you be a [opposite sex of subject] if you wanted to be?

**Story Stems**

*This test required two girl puppets, two boy puppets, and an alien puppet, illustrations of which can be found in Appendix D. This test also involved several stimulus pictures, shown in Appendix C.*

Experimenter: All right! Now I have some children to show you. This boy is named Danny [experimenter shows child the Danny puppet], this girl is named Alice [experimenter shows child the Alice puppet], this boy is named Scott [experimenter shows child the Scott puppet], this girl is named Melissa [experimenter shows child the Melissa puppet], and this alien is named Wuba [experimenter shows child the Wuba puppet, a gender neutral alien puppet associated with the first “practice” story designed to ensure the child’s understanding of the game]. I have some stories about these children, and some things that they might like to play with. In this game, I’ll show you the story, and then you can tell me what you think will happen next.
Experimenter: Let’s start with Wuba. This is Wuba [experimenter shows child the alien puppet]. Wuba goes to the kitchen in the morning. Wuba sees that there are pancakes to eat [experimenter shows the child a picture of pancakes], and there are bananas to eat [experimenter shows the child a picture of bananas]. Tell or show me what happens next.

Throughout this testing process, the experimenter allowed the child to manipulate the puppets and props. Establishing attention, patience, and curiosity on the child’s part was essential to ensure that he/she would continue active participation (Bretherton & Oppenheim, 2003). Towards this end, the experimenter made a point to demonstrate interest in the child’s completion of each of the stories employing both verbal acknowledgement and gesture; for example, the experimenter nodded as the child spoke. When the child seemed done speaking or put down a puppet, the researcher continued on with the next story stem in the same engaged and encouraging manner.

- This is Danny [experimenter shows child the Joey puppet]. Danny goes to his classroom in the morning. He sees that there is a spot to play with a truck, like this [experimenter shows the child a picture of a toy truck], and a spot to play with a baby doll, like this [experimenter shows the child a picture of a baby doll]. Tell or show me what happens next.

- This is Alice (shows girl puppet). Alice goes to her classroom in the morning. She sees that there is a spot to play with a toy bunny, like this (shows picture of toy bunny), and a spot to play with a toy hammer (shows picture of toy hammer). Tell or show me what happens next.
This is Scott (shows boy puppet). Andrew plays with his family. Andrew can help cook dinner (shows picture of food), or Andrew can play outside (shows picture of a ball). Tell or show me what happens next.

This is Melissa (shows girl puppet). Melissa plays with her dog. Melissa can give her dog a treat (shows picture of dog with treat), or Melissa can play a game with her dog (shows picture of a dog with a ball). Tell or show me what happens next.

Once each of these prompts had been presented and the activities completed, the researcher put away the pictures and got out the baskets for the next section of the game. These baskets were labeled with images—two girls and a woman, two boys and a man, or a boy, a girl, a man, and a woman—to represent who the child believed could use items in the basket. The images and pictures of items can be found in Appendix D.

For this last part of the game, I am going to show you pictures of things that people could look at, or play with, or use to work. Some of these things are more for girls and women, some of these things are more for boys and men, and some of these things are for both boys and girls and men and women. If you think the picture is mostly for girls and women, it goes in this basket. If you think the picture is mostly for boys and men, it goes in this box. If you think the picture is mostly for everyone, it goes in this basket. You can tell me which basket to put the cards in.

Practice items

Experimenter: Here is a picture of a hand. [Researcher shows picture of a hand] Is this more for girls and women, more for boys and men, or for everyone? Researcher waits for the child to respond verbally, by pointing to a basket, or by putting the picture in a basket. After the child responds, the researcher repeats their answer in a neutral tone to demonstrate interest and for purposes of coding.
• *If the child hesitates, the researcher says:* I’ll remind you about the game. Some of these things are more for girls and women, some of these things are more for boys and men, and some of these things are for both boys and girls. If you think the picture is mostly for girls, it goes in this box. If you think the picture is mostly for boys, it goes in this box. If you think the picture is mostly for everyone, it goes in this box. You can tell me which box to put the cards in. *Then the researcher repeats:* Here is a picture. Is it more for girls and women, more for boys and men, or for everyone?

• *If the child answers that they don’t know, the researcher says:* If you had to choose one, which basket would you say? More for girls and women, more for boys and men, or for everyone?

• *If the child does not put the hand in the “everyone” box, the researcher says:* Hmm, let’s try that again. This is a picture of a hand. Is it more for boys and men, more for girls and women, or for both boys and men and girls and women?

*After the child either sorted the picture correctly, or if he or she sorted the picture incorrectly once and then tried again, the researcher moved on to the next picture. This initial process was repeated for a picture of a man, which should be sorted into the “boys & men” basket, and a picture of a woman, which should be sorted into the “girls and women” basket.*

*Once the child completed the practice rounds, the researcher moved on to the test items using the above script. Like the practice items, test items were represented by pictures that were attached to 3x5 index cards; before each session was conducted, the cards were shuffled and presented to the child in a random order. Test items can be found in Appendix D.*

*Once the child seemed to understand the game (for example, sorting the picture before the researcher had finished saying all the options), the researcher proceeded by simply presenting*
each new picture and saying: Here is a picture. Throughout this process, the researcher continued to repeat verbally children’s basket choices after they had sorted each picture. 

After the child had sorted every picture: Thank you for playing my game with me! I have two stickers and you can pick one to take with you. The experimenter then walked the child back to his or her classroom.
Appendix H
Play Session Transcript

The experimenter brought two children at a time to a space outside of the classroom. This room had a table set up with toys to play birthday party, as well as a prince and princess costume including a tiara and crown. A second researcher was present in the room to video record the game. The researcher started the session by explaining the game.

Experimenter: I’m glad you are both going to play a game with me, [child’s name] and [child’s name]! My name is Riley, and this is my friend [coder’s name]. This is a game to play birthday party. At this birthday party, children have been invited to wear costumes [researcher shows the children the costumes]. You can each wear a costume and play birthday party together. My friend [coder’s name] has a video camera so we can remember how people play this game. I’m going to sit over here and watch you play. You can play birthday party however you want to!

Costumes were equally attractive prince and princess costumes. The prince costume included a crown and the princess costume included a tiara, so that children who wished to could wear only part of the costume. The researcher and coder made themselves available to help children put on the parts of the costume that they wished to wear.

- If either child objected to wearing a costume, the researcher said: This is a game where you can wear costumes. You could put on only part of the costume if you want to. If you don’t want to wear the costume you can still play birthday party!

Once children had put on costumes (or decided not to wear them), the researcher showed them where to sit at the birthday party table.

- If the children appeared to have a hard time starting to play, the researcher prompted them with any of the following phrases saying:
  - Who would go to a birthday party?
What could you do at a birthday party?

- Point out the birthday party toys if the children need more prompting

The researcher then watched the game and took note of any themes or characters that emerged.

After about 7 minutes, or when children seemed finished (whichever came first), the researcher ended the game.

Experimenter: Thank you very much for playing this game! You do have to leave the costumes and toys here so other children can have a turn to play. I have four stickers (two identical pairs of stickers) and each of you can choose one. The researcher then took the children back to their classroom.
GENDER STEREOTYPES AND PLAY

Appendix I

Observation Coding Sheet

<table>
<thead>
<tr>
<th>Child's Name:</th>
<th>Age:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Date</th>
<th>Weather</th>
<th>Indoor activities</th>
<th>Playmates</th>
<th>Type of play</th>
<th>Language</th>
<th>Roles</th>
<th>Themes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Reading</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Play-dough</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Science</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sand</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Water table</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fantasy Play</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dramatic Play</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
<td>M/F</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Action/hero play</th>
<th>Occupation feminine</th>
<th>Domestic feminine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallel play:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Children play side by side, but do not engage or acknowledge each other.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Occupational femininity</th>
<th>Domestic masculinity</th>
<th>Play themes</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animals</td>
<td>occupation masculine</td>
<td>Domestic masculine</td>
<td>Play themes</td>
</tr>
</tbody>
</table>

- Complex social pretend: Children pretend to be different characters or play out different scenarios.
- Cooperative social pretend: Children take on reciprocal roles and engage in cooperative play.
- Action/hero play: Children imagine themselves as heroes or action figures.
- Domestic feminine: Children engage in play that involves domestic roles and activities.
- Domestic masculine: Children play with toys and engage in activities that are typically associated with males.
- Animals: Children play with and interact with animal toys and figures.
- Blocks: Children build structures and play with building blocks.
- Sand: Children engage in sand play and build sand castles.
- Water table: Children play with water and engage in water-related activities.
- Reading: Children read books and engage in reading-related activities.
- Dramatic Play: Children engage in role-playing and pretend play.
- Indoor activities: Children engage in activities that take place indoors.

Type of play: 
- Solitary
- Parallel
- Simple social play
- Complementary/reciprocal
- Cooperative social pretend
- Complex social pretend

Language: 
- Words
- Sentences
- Narratives

Roles: 
- Domestic feminine
- Domestic masculine
- Animals
- Occupation feminine
- Occupation masculine

Themes: 
- Domestic feminine
- Domestic masculine
- Animals
- Action/hero play

Date

Weather

Indoor activities

Playmates

Type of play

Language

Roles

Themes